Advances in Prevention Science

Zili Sloboda
Hanno Petras *Editors*

Defining Prevention Science

Springer

Advances in Prevention Science

For further volumes:
http://www.springer.com/series/8822

Zili Sloboda • Hanno Petras
Editors

Defining Prevention Science

Springer

Editors
Zili Sloboda
Ontario, Ohio
USA

Hanno Petras
JBS International, Inc.
Rockville, Maryland
USA

ISBN 978-1-4899-7423-5 ISBN 978-1-4899-7424-2 (eBook)
DOI 10.1007/978-1-4899-7424-2
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014931239

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Prevention Science: A Global Issue

There are few countries or populations which are not affected by the consequences of behavioural risk (Ezzati et al., 2002). Obesity, for example, was once considered primarily a problem of high-income countries, but its consequences are now linked to more deaths worldwide than a lack of food (WHO, 2004). Despite young people traditionally being perceived as more healthy than older members of the population, a large proportion of premature adult deaths (up to 70 %) are associated with behaviours initiated in adolescence (Resnick, Catalano, Sawyer, Viner, & Patton, 2012), and young people's involvement in preventable risks such as drug and alcohol use or engagement in unsafe sex are now important contributors to the global burden of disease (Gore et al., 2011). The consequences of these behaviours in both adults and young people are also expensive. Although it is difficult to estimate the costs placed on health and social services as a result of preventable disease, good population health is considered a prerequisite of current economic growth strategies (e.g. Health 2020 in Europe). This global priority is also echoed in the World Health Organisation's *Action Plan for Global Strategy for the Prevention and Control of Noncommunicable Diseases* (2008–2013) (WHO, 2008), which outlines activities to prevent and control the four main non-communicable diseases; cardiovascular disease, diabetes, cancers and chronic respiratory diseases. To help achieve this, the Action Plan recommends developing and promoting prevention interventions that are designed to reduce shared risk factors (tobacco use, unhealthy diets, physical inactivity and harmful use of alcohol) (WHO, 2008). Underpinning this priority is the recognition that success is dependent upon international collaboration and shared learning; indeed the current economic climate means that sharing knowledge and resources is the only viable forward for many countries. However, even operating within an international framework, the challenge for the prevention field is to develop and disseminate evidence-based policies that not only ensure equality of access to interventions but are also sensitive to local moderators (Resnick et al., 2012).

Great advances have been made in recent years, in diverse fields ranging from molecular biology to behavioural epidemiology, that have allowed for the identification of important psychobiological, developmental and environmental mediators of

risk behaviour (Catalano et al., 2012). Many of these are common across cultures, but equally, many are not (e.g. Beyers, Toumbourou, Catalano, Arthur, & Hawkins, 2004). Indeed, the influence of socioeconomic status and social inequalities on health and well-being, for example, can differ even within small geographies (Marmot Review, 2010). This means that global prevention success depends not only upon robust theory and demonstration of programme efficacy but also upon consideration of the implementation, transferability and adaptation of programmes across diverse delivery systems and policy objectives (Brotherhood & Sumnall, 2011). Interventions that have demonstrated effectiveness in their country of development (most commonly the USA) are not easily translated, and notable examples exist of beneficial programme effects that are not sustained when transferred internationally (Fraser et al., 2011; Malti, Ribeaud, & Eisner, 2011), especially when assessed by independent research teams (Eisner, 2009). This not only suggests a need for theory-driven adaptation processes (Ferrer-Wreder, Sundell, & Mansoory, 2012) but also highlights the importance of independent international replication of intervention effects.

Similarly, moving away from intervention theory towards implementation, we must also consider the diverse range of sociocultural environments into which prevention is likely to be delivered (Li, Mattes, McMurray, Hertzmanm, & Stanley, 2009). Interventions, particularly those targeted towards young people may be most effective when delivered in accordance with (or are at least sensitive to) prevailing social trends and attitudes (Room, 2012). However, because of their diversity, these may often be in opposition to both international policies and conventions (e.g. UN Single Convention on Narcotic Drugs), and the priorities of health and social care professionals (e.g. prevention of cannabis use). These differences can be seen reflected in international epidemiological data on preventable behaviours (e.g. Hibbell et al., 2012), public preferences towards regulation of risk (Gallup Organisation, 2011), political, social and cultural norms on acceptability of risky (Nutt, 2009), and cross-national differences in policy regulating access to risky behaviours (e.g. Joosen & Raw, 2006). For example, there is a sixfold difference in per capita consumption of pure alcohol between the lowest and highest consuming countries (WHO, 2011), and global differences in alcohol policy manifest in an array of marketing regulations (e.g. sports sponsorship bans), interventionist pricing policies (e.g. minimum alcohol unit pricing), cultural norms (e.g. legal alcohol purchase age), and the acceptability of industry lobbying (such as industry self-regulation and "responsibility deals") (Babor et al., 2003). Subsequently, what might be considered an acceptable prevention strategy in one country (e.g. "dry" University campuses) may be rejected by another. Similarly, prevention success as defined in one country (e.g. alcohol abstention as the goal of prevention) might not be acceptable in another (e.g. where an alcohol harm reduction approach is preferred).

It is into such complex mixes of behavioural moderators that we try to introduce prevention work, and so it is unsurprising that our findings are often not transferable. As prevention researchers we can have faith in the robustness of our theories, and can modify our interventions in relation to specific and consistent population characteristics. However, we often assume populations welcome intervention, but recipients of prevention are dynamic social actors who face multiple competing and increasingly

global influences on their behaviour (Labonte, Mohindra, & Schrecker, 2011). The introduction of tobacco control policy, for example, has been a relative public health success in the developed world, but it has also shown us that as one market is restricted, others are exploited (Otañez, Mamudu, & Glantz, 2009). Most prevention researchers are unable to influence such macro determinants of health and well-being, but as challenged by the World Health Organisation, it is through international collaboration that we might begin to address the major global health and social challenges we face. This is of course easier to write than enact, and successful action will need the cooperation of many stakeholders, from the general public to international organisations. By taking a global perspective on prevention, particularly through our activities in international fora, we might more readily achieve this. This first volume in the Advances in Prevention Science Book Series is also to be welcomed as it will help to create a solid foundation for international prevention science activities. It provides a universal prevention lexicon, and outlines evidence-based theories and methods that will support a unified approach to preventing and managing engagement in unhealthy and risky behaviours across the lifespan.

Harry Sumnall
Centre for Public Health
Liverpool John Moores University
Liverpool, UK

EU-Society for Prevention Research
Warsaw, Poland

References

Babor, T. F., Caetano, R., Casswell, S., Edwards, G., Giesbrecht, N., Graham, K., et al. (2003). *Alcohol: No ordinary commodity—Research and public policy*. Oxford: Oxford University Press.

Beyers, J. M., Toumbourou, J. W., Catalano, R. F., Arthur, M. W., & Hawkins, J. D. (2004). A cross-national comparison of risk and protective factors for adolescent substance use: The United States and Australia. *Journal of Adolescent Health, 35*, 3–16.

Brotherhood, A., & Sumnall, H. R. (2011). *European drug prevention quality standards*. Lisbon: EMCDDA.

Catalano, R. F., Fagan, A. A., Gavin, L. E., Greenberg, M. T., Irwin, C. E., Ross, D. A., et al. (2012). Worldwide application of prevention science in adolescent health. *Lancet, 379*, 1653–1664.

Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology, 5*, 163–183.

Ezzati, M., Lopez, A. D., Rodgers, A., Vander Hoorn, S., Murray, C. J., & Comparative Risk Assessment Collaborating Group. (2002). Selected major risk factors and global and regional burden of disease. *Lancet, 360*, 1347–1360.

Ferrer-Wreder, L., Sundell, K., & Mansoory, S. (2012). Tinkering with perfection: Theory development in the intervention cultural adaptation field. *Child & Youth Care Forum, 41*, 149–171.

Fraser, M. W., Guo, S., Ellis, A. R., Thompson, A. M., Wike, T. L., & Li, J. (2011). Outcome studies of social, behavioral, and educational interventions emerging issues and challenges. *Research on Social Work Practice, 21*, 619–635.

Gallup Organisation. (2011). *Flash Eurobarometer 330: Youth attitudes on drugs*. Online publication. Retrieved on June 25, 2012, from http://ec.europa.eu/public_opinion/flash/fl_330_en.pdf

Gore, F. M., Bloem, P. J. N., Patton, G. C., Ferguson, J., Joseph, V., Coffey, C., et al. (2011). Global burden of disease in young people aged 10–24 years: A systematic analysis. *Lancet, 377*, 2093–2102.

Hibbell, B., Guttormsson, U., Ahlström, S., Balakireva, O., Bjarnason, T., Kokkevi, A., et al. (2012). *The 2011 ESPAD report - Substance use among students in 36 European countries*. Stockholm: CAN.

Joosen, L., & Raw, M. (2006). The tobacco control scale: A new scale to measure country activity. *Tobacco Control, 15*, 247–253.

Labonte, R., Mohindra, K., & Schrecker, T. (2011). The growing impact of globalization for health and public health practice. *Annual Review of Public Health, 32*, 263–283.

Li, J., Mattes, E., McMurray, A., Hertzmanm, C., & Stanley, F. (2009). Social determinants of child health and wellbeing. *Health Sociology Review, 18*, 3–11.

Malti, T., Ribeaud, D., & Eisner, M. (2011). The effects of two universal preventive interventions to reduce children's externalizing behavior: A cluster randomized controlled trial. *Journal of Clinical Child and Adolescent Psychology, 40*, 677–692.

Marmot Review. (2010). *Fair society, healthy lives: Strategic review of health inequalities in England Post 2010*. London: Marmot Review.

Nutt, D. J. (2009). Equasy – An overlooked addiction with implications for the current debate on drug harms. *Journal of Psychopharmacology, 23*, 3–5.

Otañez, M. G., Mamudu, H. M., & Glantz, S. A. (2009). Tobacco companies' use of developing countries' economic reliance on tobacco to lobby against global tobacco control: The case of Malawi. *American Journal of Public Health, 99*, 1759–1771.

Resnick, M. D., Catalano, R. F., Sawyer, S. M., Viner, R., & Patton, G. C. (2012). Seizing the opportunity of adolescent health. *Lancet, 379*, 1564–1567.

Room, R. (2012). Preventing youthful substance use and harm. Between effectiveness and political wishfulness. *Substance Use and Misuse, 47*, 936–943.

World Health Organisation (WHO). (2004). *Global strategy on diet, physical activity, and health*. Geneva: World Health Organisation.

World Health Organisation (WHO). (2008). *2008–2013 Action plan for the global strategy for the prevention and control of noncommunicable diseases*. Geneva: World Health Organisation.

World Health Organisation (WHO). (2011). *Global status report on alcohol and health 2011*. Geneva: World Health Organisation.

The Promise of Prevention Science

Prevention Science is a relatively new field, with some of the first papers outlining the conceptual framework, methods and research priorities having been published in the early 1990s (e.g., Coie et al., 1993; Kellam & Van Horn, 1997). These early descriptions highlighted the breadth of prevention science, which draws from multiple disciplines and encompasses a broad range of research including studies of epidemiology, studies designed to identify risk and protective factors of a problem or disorder and the development of interventions for preventing or ameliorating high-risk behaviours, disease, disorder or injury. In addition, prevention science also includes research on the translation and dissemination of effective preventive interventions into practice; studies to understand the science of bringing efficacious and effective interventions to scale in order to have public health impact.

Significant progress in prevention has been made over the last 30 years. Epidemiological and etiological research has identified numerous biological, psychological, social and environmental risk and protective factors that influence behaviour and positive well-being as well as disorders and illness. This research has contributed to the development of programs and policies that have demonstrated efficacy to prevent behavioural and health problems and promote well-being by targeting these empirically identified risk and protective factors, representing great potential for enhancing public health and well-being. When carefully implemented, such interventions can prevent a wide range of health problems, promote positive development and achieve economic benefits (CDC, 2007; National Prevention Strategy, 2011). Considerable evaluation data indicate that these types of interventions have had significant and far-reaching effects in reducing unhealthy eating; physical inactivity; alcohol, tobacco and other drug abuse; teen pregnancy; school failure; delinquent behaviour; violence; and other mental, emotional, behavioural and physical health problems. Furthermore, these interventions have shown a cost-beneficial economic impact on education, criminal justice, social and health services (O'Connell, Boat, & Warner, 2009).

Based on these data and on recognition of this potential, a number of current federal initiatives have been put in place to support scale-up of evidence-based programs (e.g., home-visitation, teen pregnancy, social innovation fund). Several of these efforts include a tiered approach to funding, with greater funding available for

those programs that meet the highest level of evidentiary standards (i.e., typically randomized controlled trials), making additional resources available for evaluating programs with some supportive evidence of effects, and encouraging agencies to innovate and test ideas with strong potential that are supported by preliminary research findings. These initiatives have been met with tremendous enthusiasm, both because of the potential to positively impact our nations' health and decrease risk for a range of social problems, and also because of the ability to advance knowledge regarding what works and how evidence-based preventive interventions can be effectively implemented and evaluated at scale.

Despite the impressive progress made, considerably more work is needed to advance both the science and practice of prevention. Research is needed to test new approaches, identify mediating mechanisms of intervention effects, understand factors associated with poor implementation fidelity and how to surmount them, determine whether adaptation is necessary to make prevention programs suitable to different populations as well as how to adapt evidence-based prevention programs without undermining their effectiveness. Research is needed to test the optimal combination of prevention programs together to create effective comprehensive prevention strategies and extend current prevention findings to multiple problem behaviours.

This volume is important in aiding in these efforts, as it is one of the first to provide a comprehensive set of chapters devoted to theory, research design and analytic techniques specific to prevention science. Mainly, the editors are intentional in their effort to bring a global perspective to the field, including chapters written by an international set of experts. Increasing collaboration will allow for greater flow of information across geographical boundaries and provides opportunities to apply principles and programs in multiple contexts—to better understand questions of how, for whom and under what circumstances.

In June, 2010 President Obama signed an Executive Order creating the National Prevention, Health Promotion, and Public Health Council. Charged with creating a National Prevention Strategy, the formation of the Council and the focus on an integrated national strategy provided an unprecedented opportunity to shift the United States from a focus on sickness and disease to one based on wellness and prevention (National Prevention Strategy, 2011). The goals of the Council are ambitious and largely based in research that has been conducted in prevention science. By continuing to advance the science of prevention, there is tremendous potential for improving not only this but also all nations' health by reducing the incidence of morbidity, mortality and the associated social costs of problems and disorders such as drug addiction and alcoholism, HIV/AIDS, mental illness, cancer, cardiovascular disease, obesity and violence. This volume provides a solid foundation to aid in these goals.

Chicago, IL

Deborah Gorman-Smith
Society for Prevention Research
University of Chicago

Centers for Disease Control and Prevention. (2007). *Improving public health practice through translation research (R18)*. Retrieved from, http://grants.nih.gov/grants/guide/rfa-files/RFA-CD-07-005.html

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., et al. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist, 48*, 1013–1022.

Kellam, S., & Van Horn, Y. V. (1997). Life course development, community epidemiology and preventive trials: A scientific structure for prevention research. *American Journal of Community Psychology, 25*, 177–188.

National Prevention, Health Promotion, and Public Health Council. (2011). *National prevention strategy*. Washington, DC: National Prevention, Health Promotion, and Public Health Council. Retrieved from, www.healthcare.gov/center/councils/nphpphc/strategy/report.pdf

National Prevention Strategy. (2011). Retrieved from, http://www.surgeongeneral.gov/initiatives/prevention/strategy/report.pdf

National Research Council and Institute of Medicine. (2009). Preventing mental, emotional, and behavioural disorders among young people: Progress and possibilities. In Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth, and Young Adults, M. E. O'Connell, T. Boat, & K. E. Warner (Eds.), *Research advances and promising interventions*. Washington, DC: The National Academies Press.

Preface

To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science.

Albert Einstein

The maturity of an evolving science is marked by the development of its own theories, ethics, methodologies and lexicon. Prevention research has progressed over the past three decades towards becoming the basis for a new science, prevention science, defined as: a science that tests theory-based hypotheses towards enhancing or refining the process that makes individuals' engagement in healthy and non-risky behaviours more likely (Biglan et al., 2011).

The establishment of the American Society for Prevention Research (SPR) in 1991 and the European Society for Prevention Research (EU-SPR) in 2010 has increased the recognition that the prevention of mental, emotional and behavioural problems shares similar approaches, requires similar research designs and warrants its own terminology, skills and competencies. In response, in 2011, SPR published Standards of Knowledge for the Science of Prevention (http://www.preventionresearch.org/Society%20for%20Prevention%20Research%20Standards%20of%20Knowledge.pdf) that laid out the foundation for this new field, drawing from epidemiology, psychology, medicine, sociology, social work, education, economics, public health, biostatistics, geography, anthropology, policy analysis, criminology, neuroscience and genetics. The work lays out three domains, Epidemiology, Intervention Development, and Research Methodology (to include design and statistical applications) and the prevention-related competencies of each.

Taking the Standards of Knowledge document as a point of departure, the focus of this book lies in documenting the state of the art in prevention knowledge while at the same time emphasizing the importance of developmental benchmarks and competencies as a basis for understanding vulnerability, socialisation and decision making. In addition, the book integrates implementation and prevention science, and, presents cutting-edge concepts related to alternative research designs for evaluations of prevention interventions and the integration of cost assessment into prevention program evaluations.

As with any other science, we believe that Prevention Science should not be confined by geographic or cultural boundaries. To start the international dialogue, we therefore wanted to include authors from outside the United States and were pleased by the enthusiastic response of our European colleagues, many of whom are members of the EU-SPR. We wish to acknowledge with great appreciation the contributors to this first book on prevention science. Their names and affiliations are listed below. We want to thank them for sharing our vision and especially for their patience with the process of creating the book and their tolerance in addressing our comments and edits.

We also wish to thank Wendy Caron of JBS International, Inc. for her assistance in the formatting and editing of the chapters. Her attention to details in the text and references is much appreciated.

We particularly wish to thank Gail Bassin and Jerri Shaw, the CEO and President of JBS International, Inc., for allowing us the time to work on this book. Without their enthusiastic support, the book most likely would have not come to fruition.

Finally, we would like to thank our Springer editor, Khristine Queja. She has patiently worked with us over the past 3 years to develop not only this first book but three others in the series, Preventing Substance Use, Prevention Science in School Settings, and, Preventing Crime and Violence.

The final chapter in this book by Dr. Fabrizio Faggiano and his colleagues Fabrizia Giannotta and Elias Allara provide recommendations for the next advancements in prevention science. This chapter sets out an agenda and a challenge for prevention scientists both those who conduct the research and also those who deliver the prevention interventions. We all want to assure the delivery of evidence-based prevention programming to improve and enhance the lives of ourselves and future generations.

We view this book not as the final word on prevention science but as a starting point that will encourage discourse, discussion and refinement. In line with the quote by Albert Einstein, we hope that this book can help to initiate the creative process of defining and refining Prevention Science.

Ontario, OH — Zili Sloboda
Rockville, MD — Hanno Petras

Biglan, A., Domitrovich, C., Ernst, J., Etz, K., Mason, M. J., Robertson, E., et al. (2011). *Standards of knowledge*. Fairfax, VA: Society for Prevention Research.

Contents

List of Contributors

Katherine M. Aguirre Department of Psychology, University of Texas, El Paso, TX, USA

Elias Allara School of Public Health, University of Torino, Torino, Italy

James C. (Jim) Anthony Epidemiology & Biostatistics, Michigan State University, East Lansing, MI, USA

Karl E. Bergmann Department for the Epidemiology of Non-communicable Diseases, Robert Koch-Institute, Berlin, Germany

Shelley A. Blozis Department of Psychology, University of California, Davis, CA, USA

C. Hendricks Brown Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA

Eric C. Brown Social Development Research Group, School of Social Work, University of Washington, Seattle, WA, USA

Felipe González Castro Department of Psychology, University of Texas, El Paso, TX, USA

Jonathan P. Caulkins Heinz College, Carnegie Mellon University, Pittsburgh, PA, USA

H. John Heinz III School of Public Policy and Management, Pittsburgh, PA, USA

Rand D. Conger Family Research Group, University of California, Davis, CA, USA

David Cordova School of Social Work, University of Michigan, Ann Arbor, MI, USA

Sokratis Dinos Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Linda Dusenbury Collaborative for Academic, Social, and Emotional Learning (CASEL), Chicago, IL, USA

Tanglewood Research, Inc., Greensboro, NC, USA

J. Mark Eddy Partners for Our Children, University of Washington, Seattle, WA, USA

Yannine Estrada Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA

Fabrizio Faggiano Department of Translational Medicine, Avogadro University of Eastern Piedmont, Novara, Italy

Amanda J. Fairchild Department of Psychology, University of South Carolina, Barnwell College, Columbia, SC, USA

Julee P. Farley Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Mark L. Garibaldi American Institutes for Research, Washington, DC, USA

Fabrizia Giannotta Center for Developmental Research, Orebro University, Orebro, Sweden

Deborah Gorman-Smith School of Social Administration, University of Chicago, Chicago, IL, USA

Society for Prevention Research, Fairfax, VA, USA

Julia A. Graber Department of Psychology, University of Florida, Gainesville, FL, USA

William B. Hansen Tanglewood Research, Inc., Greensboro, NC, USA

Shelley R. Hart Department of Education, Gevirtz Graduate School of Education, University of California, Santa Barbara, CA, USA

Richard E. Heyman Cardiology and Comprehensive Care, College of Dentistry, New York University, New York, NY, USA

Julie C. Hill Department of Psychology, University of Florida, Gainesville, FL, USA

Shi Huang Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA

Johan Jongbloet Department of Education, Health and Social Work, University College Ghent, Ghent, Belgium

Joshua G. Kellison Department of Psychology, Arizona State University, Tempe, AZ, USA

Kimberly Kendziora American Institutes for Research, Washington, DC, USA

Jungmeen Kim-Spoon Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Alexandra M. Kueider Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

David Levy Department of Oncology, Georgetown University, Washington, DC, USA

David P. MacKinnon Department of Psychology, Arizona State University, Tempe, AZ, USA

Shandey Malcolm Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA

Charles R. Martinez Jr. College of Education, University of Oregon, Eugene, OR, USA

Michael Mason Commonwealth Institute on Child and Family Studies, Virginia Commonwealth University, Richmond, VA, USA

W. Alex Mason National Research Institute for Child and Family Studies, Boys Town, NE, USA

Katherine E. Masyn Harvard Graduate School of Education Cambridge, Harvard University, Cambridge, MA, USA

Carol W. Metzler Oregon Research Institute, Eugene, OR, USA

Osvaldo F. Morera Department of Psychology, University of Texas, El Paso, TX, USA

Nancy Nicosia RAND Drug Policy Research Center, RAND Corporation, Boston, MA, USA

Karen Nylund-Gibson Department of Education, Gevirtz Graduate School of Education, University of California, Santa Barbara, CA, USA

David Osher Health and Social Development Program, American Institutes for Research (AIR), Washington, DC, USA

Rosalie Liccardo Pacula RAND Drug Policy Research Center, RAND Corporation, Santa Monica, CA, USA

Hilda Pantin Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA

Jeanine M. Parisi Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Hanno Petras JBS International, Inc., North Bethesda, MD, USA

Guillermo Prado Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA

Joel Raffel Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Hellgard Rauh Department of Psychologie, Universität Potsdam, Potsdam, Germany

George W. Rebok Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Louise A. Rohrbach Preventive Medicine, University of Southern California, Los Angeles, CA, USA

Mary E. Saczawa Department of Psychology, University of Miami, Gainesville, FL, USA

Zili Sloboda Ontario, OH, USA

Elizabeth Spier American Institute for Research, Washington, DC, USA

Stephen Stansfeld Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Harry Sumnall Centre for Public Health, Liverpool Johns Moores University, Liverpool, UK

EU-Society for Prevention Research, Warsaw, Poland

Emily E. Tanner-Smith Peabody Research Institute, Vanderbilt University, Nashville, TN, USA

Zoe E. Taylor Department of Human Development and Family Studies, Purdue University, West Lafayette, IN, USA

Peer van der Kreeft Department of Education, Health and Social Work, University College Ghent, Ghent, Belgium

Tina Van Havere Department of Education, Health and Social Work, University College Ghent, Ghent, Belgium

Sandra Jo Wilson Peabody Research Institute, Vanderbilt University, Nashville, TN, USA

Chapter 1
Prevention Science: An Epidemiological Approach

David Cordova, Yannine Estrada, Shandey Malcolm, Shi Huang, C. Hendricks Brown, Hilda Pantin, and Guillermo Prado

Introduction

This chapter provides an overview of prevention science and the role of epidemiology in the field of prevention science. Specifically, we discuss several major ways in which prevention science is informed by epidemiology. First, we describe how epidemiology is useful in identifying target populations and vulnerable periods as well as discuss the distribution of disease, etiological risk and protective factors, and human development across the lifespan. Second, we highlight the ways in which epidemiology is used to develop frameworks, including an ecological and developmental framework to prevention, and how these frameworks are useful for understanding and comparing targeted populations for specific times of risk. Third, we describe experimental epidemiology and widely used analytic and methodological approaches for testing the efficacy and effectiveness of interventions. Fourth, we describe the role of epidemiology in implementation strategies. Finally, we discuss the need to work toward feedback loops whereby prevention science findings are used to inform epidemiology and vice versa. Although the fields of prevention science and epidemiology share a common goal and interest in health promotion and disease prevention, to some extent, both disciplines often operate in isolation with a minimal feedback loop process. These feedback loops may be essential in advancing both fields. To highlight these concepts, we use several disease and health-risk behaviors, including HIV/AIDS, obesity, and alcohol and drug use.

D. Cordova (✉)
School of Social Work, University of Michigan, 1080 S. University, Suite 2846, Ann Arbor, MI 48109, USA
e-mail: cordovad@umich.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_1,

Introduction to Prevention Science and Epidemiology

Prevention science is concerned with, among other things, identifying antecedents that impact health and health behavior, as well as the development of models to ameliorate undesired health outcomes and promote health behavior. A relatively new field, prevention science is interdisciplinary and combines life-course development, community epidemiology, and preventive intervention perspectives. Epidemiology in particular plays an important role in prevention science. Below, we describe the role of epidemiology in the field of prevention science.

Traditionally, epidemiology has been defined as the study of the distribution and determinants of health; its aim is to prevent and reduce, survey, and control health disorders (Susser & Stein, 2009). As the cornerstone method of public health research, epidemiology plays a key role in prevention science. Of particular importance to the field of epidemiology is describing the natural history or career of the outcome of interest. The natural history can be defined as the progression of the outcome of interest from the time of exposure to cessation/desistance (Gordis, 2009). From this perspective, central to epidemiology is to identify those factors that contribute to the onset, progression/escalation, and cessation/desistance of the behavior or problem of interest. Using substance use as an example, epidemiology is interested in identifying factors contributing to the onset of substance use, risk and protective factors, whom to target and at what developmental stage, and the methods and design that may be most optimal for highlighting the natural history of substance use. For example, epidemiologic studies show us that early onset of substance use among adolescents, an important period of development, is associated with a greater likelihood of later abuse and dependence (Behrendt, Wittchen, Höfler, Lieb, & Beesdo, 2009; Grant & Dawson, 1998), as well as identify those correlates and risk factors that may contribute to the escalation or progression from use to abuse and dependence (Swendsen et al., 2012).

Epidemiological concepts, including the distribution of disease, determinants of health, and the role of host-environment-agent in understanding disease and etiology, are important in advancing the field of prevention science. These concepts are described below.

The distribution of disease refers to the frequency, pattern, and history of a particular condition, disorder, or disease among groups or populations. Identifying the distribution of disease is helpful in determining segments of the population that are most affected by a disease, condition, or disorder and plays an important role in informing the development of preventive interventions for specific subpopulations, including racial/ethnic minorities, developmental age groups, and/or those who live in certain geographic regions. For example, epidemiologic research indicates that, when compared with their non-Hispanic White and African American counterparts, Hispanic youth report the highest lifetime, annual, and 30-day prevalence rates of both licit and illicit substance use, excluding amphetamines (Johnston, O'Malley, Bachman, & Schulenberg, 2013). These epidemiologic research findings indicate that specific populations may require specialized prevention services, including the

development of Hispanic-specific preventive interventions (Prado et al., 2007, 2012). The distribution of disease, however, does not operate in isolation and is influenced by genetic, environmental, and social determinants of health that also involve risk and protective factors (Susser & Stein, 2009).

The determinants of health can be defined as those risk and protective factors that may have an impact on the distribution of disease and occurrence of a condition or an event (Torrence, 1997). Models of causation, including the host-environment-agent, are useful for better understanding the role of risk and protective factors associated with physical and psychological health disorders.

The host-environment-agent model is used to describe the intersection of the host (e.g., an individual), the environment (e.g., vehicle), and agent (e.g., driver distraction) that interact with one another in the development of a condition such as motor vehicle injuries. The host is the individual and his/her inherent characteristics (e.g., genetic, psychological), which may be a precursor for the development of a particular health or psychological condition. Across different prevention fields, the role of agent can vary and refers to the organism or direct cause of the condition. Using motor vehicle injuries and driver distraction as an example, the agent could include the use of mobile communication devices. The environment includes all external factors that may contribute to the development of a condition and interacts with an individual's susceptibility. For example, accessibility of mobile communication devices, cultural norms and practices, and legislation on the use of mobile communication devices while driving are environmental factors that may impact a driver's use of mobile communication devices and consequently motor vehicle injuries and through which risk and protective factors can be targeted by preventive interventions. In fact, prevention science findings have highlighted the significant consequences and public health concern of using mobile communication devices while driving and motor vehicle injuries (Ibrahim, Anderson, Burris, & Wagenaar, 2011). For example, findings from the 2008 *Fatality Analysis Reporting System* (National Highway Traffic Safety Administration, 2009) indicate that approximately 16 % of all reported fatal crashes are attributable to driver distraction, including the use of mobile communication devices. Although it is still unclear how restricting the use of mobile communication devices while driving impacts motor vehicle injuries, these data have led to the implementation of preventive legislation (i.e., environment) on the use of communication mobile devices (i.e., agent) while driving (i.e., individual, host) in 39 states and the District of Colombia (Ibrahim et al., 2011). Thus, the host-environment-agent model has helped guide the field of prevention science and the development and evaluation of motor vehicle injury legislation prevention models to target risk and protective factors within the host, environment, and agent domains (Ibrahim et al., 2011). Therefore, etiological frameworks, including the host-environment-agent model, facilitate advances in the field of prevention science by providing a conceptual framework for better understanding problems, diseases, and disorders; selecting settings and stages of life that may be particularly conducive to intervention; and guiding the choice of what constructs to measure.

To summarize, epidemiology aims to contribute to the prevention of health disorders and health-risk behaviors by providing models through which the natural

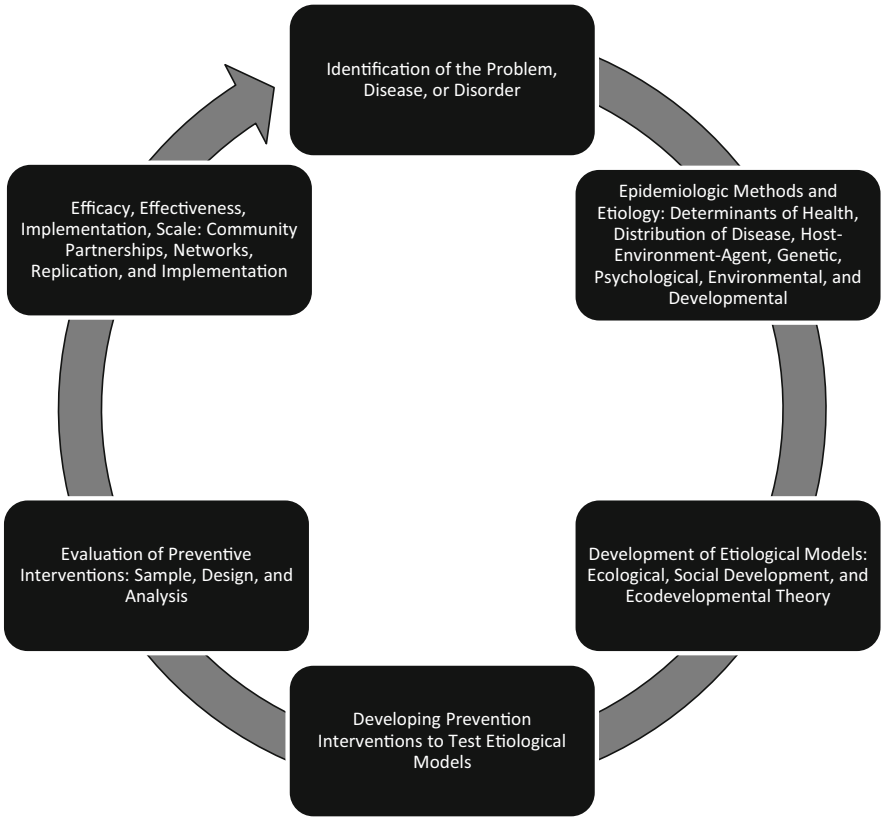

Fig. 1.1 The role of epidemiology in prevention science. Adapted from Sloboda, Glantz, and Tarter (2012)

history of disease, along with the potential risk and protective factors associated with it, can be described and defined (Torrence, 1997). As shown in Fig. 1.1, the roles of the distribution of disease, determinants of health, and the host-environment-agent are key first steps in understanding disease and etiology, as well as informing the development of prevention programs. In the next section, we describe the role of etiological models in the development of preventive interventions.

Etiology

A relatively new perspective has emerged shifting focus to the developmental vulnerability of individuals in context. This paradigm shift allows us to better understand the etiology of the outcome of interest, as well as who will and who

will not engage in certain risky and unhealthy behaviors (Sloboda et al., 2012). Specifically, the etiology of health-risk behaviors, such as adolescent drug abuse, may best be understood through a risk and protective factor model, informed by a developmental framework, that examines the intersection of genetic, psychological, and environmental factors (Prado et al., 2009; Sloboda et al., 2012). As early as the mid-1960s, researchers began to identify a host of risk and protective factors associated with behavioral health problems (Berrueta-Clement, Schweinhart, Barnett, & Weikart, 1987). Identified risk and protective factors include individual-level (e.g., genetic, psychological), proximal environmental (e.g., family, school, peers, and work), and distal environmental factors such as community and social/economic conditions (Kellam & Langevin, 2003). Furthermore, researchers state that risk and protective factors do not operate in isolation and should be studied as an integrated developmental process that also includes genetic factors (Schwartz, Pantin, Coatsworth, & Szapocznik, 2007). Risk and protective factor models have great utility to the field of prevention science, particularly as it relates to identifying different points of intervention that prevention programs may target.

Genetics

Research has highlighted the ways in which genetic markers play an instrumental role in the predisposition to engage in, for example, alcohol and drug use. In fact, studies examining heritability estimates suggest that genetics account for as much as 40–70 % of the risk and susceptibility in the development of alcohol or drug abuse and dependence (Heath et al., 1997). Thus, prevention science may play an important role in ameliorating health-risk behaviors, including alcohol or drug abuse and dependence, particularly through a better understanding of the process of genetic markers. For example, it is hypothesized that adolescents with one or two copies of the short variant 5HTTLPR genetic marker may be more susceptible to engaging in health-risk behaviors, relative to adolescents with two copies of the long variant 5HTTLPR genetic marker. Prevention scientists have demonstrated that participating in a preventive intervention may mitigate the risk of the 5HTTLPR genetic marker, thereby ameliorating the genetic predisposition for engaging in health-risk behaviors (Brody, Beach, Philibert, Chen, & Murry, 2009; Brody, Chen, Beach, Philibert, & Kogan, 2009). Therefore, disease and problem behaviors are best understood in the context of the interplay between genetics and environment. For example, gene × environment approaches have played a significant role in advancing the field of prevention science, including working toward a better understanding of genetic susceptibility and cigarette smoking in people with Crohn's disease (Helbig et al., 2012) and gene × lifestyle and gene × drug interactions and obesity (Franks & Poveda, 2011). Advances in technology, methodology, and statistics will further our understanding of the role of gene × environment in prevention science, as well as advance rodent models to human models to

identify the disease or problem behavior's susceptibility pathways (Blackburn & Jerry, 2011). The gene × environment model, therefore, is very promising with respect to advancing the etiology literature, which in turn has the potential to maximize the efficacy and effectiveness of prevention programs (Howe, Beach, & Brody, 2010).

Psychological Factors

In addition to genetic factors, psychological or intrapersonal factors are important in understanding health-risk behaviors. Psychological or intrapersonal factors refer to those cognitive, emotional, and attitudinal processes that vary across individuals and can serve as risk or protection for health and behavior (Hemphill et al., 2011; Tonin, Burrow-Sanchez, Harrison, & Kircher, 2008). For example, adolescents who report positive attitudes toward alcohol or drug use may be more vulnerable or susceptible to engaging in alcohol or drug use compared with youth who report more negative attitudes toward alcohol or drug use (Cordova et al., 2011; Hemphill et al., 2011; Tonin et al., 2008). To this end, prevention science may play an important role in ameliorating negative psychological or intrapersonal factors. For example, cognitive and attitudinal processes are pathways through which prevention programs have demonstrated efficacy in preventing/reducing alcohol and drug use among adolescents (Hops et al., 2011).

Environmental/Ecological Factors

Environmental/ecological factors also play a role in the development of health conditions. Environmental/ecological factors are those social, cultural, and contextual processes, including family, peer, community, and legislation, that both influence and are influenced by the individual (Bronfenbrenner, 1979, 1989; Catalano & Hawkins, 1996; Szapocznik & Coatsworth, 1999). A substantial amount of research has demonstrated the ways in which environmental/ecological factors influence risk or protection for health and mental health (Catalano & Hawkins, 1996; Cleveland, Feinberg, & Greenberg, 2010; Griffin & Botvin, 2010). For example, studies suggest that communities characterized by higher levels of social capital and neighborhood collective efficacy may have protective effects on depression among certain Hispanic subgroups (Vega, Ang, Rodriguez, & Finch, 2011). Thus, the environmental/ecological context is important in better understanding risk and protection for health and mental health. Prevention programs that target environmental/ecological processes at multiple contextual levels may be particularly helpful in decreasing health-risk behaviors and promoting health (Hawkins et al., 2012). In fact, Hawkins and colleagues (2012) have shown that, when compared with youth in control communities, youth in Communities That Care, a community-based prevention

program, demonstrate a decrease in community-wide levels of health-risk behaviors, including substance use over time.

Etiology Across the Lifespan

The role of genetics, psychological, and environmental/ecological factors on health and health-risk behaviors should be informed through a developmental and life-course lens (Kellam et al., 1991; Sloboda et al., 2012). For example, we know that individuals might be at increased risk for alcohol and drug use in certain developmental periods across the human development life cycle. In fact, research has shown that adolescence and early adulthood in particular are developmental stages in which drug abuse might be more pronounced and likely to occur, compared with middle adulthood (Botvin, Griffin, Paul, & Macaulay, 2003; Dishion, Kavanagh, Schneiger, Nelson, & Kaufman, 2002; Hawkins, Catalano, & Miller, 1992). Therefore, adopting a developmental and life-course perspective to inform prevention science might be important in interrupting the sequelae of negative health and psychological outcomes during critical periods across the lifespan (Braveman & Barclay, 2009).

Developmental Epidemiology

Informed by both a developmental and life-course perspective, developmental epidemiology identifies specific proximal individual or environmental factors at an early stage of life to then target preventive interventions at these factors (Kellam & Langevin, 2003). Interventions are then evaluated to determine whether and the extent to which targeting the identified risk factors has positively impacted more distal factors across both time and development. Critical periods are an important concept in developmental epidemiology and are essential to describing the individual's life course (Braveman & Barclay, 2009). A critical period refers to "a window of time during the life course when a given exposure has a critical or even permanent influence on later health" (Braveman & Barclay, 2009, p. S164). Therefore, a critical period refers to those critical benchmarks, including age, gender, and cultural relevance and expectations that might interact during certain times of transitions across the life course (Chambers, Taylor, & Potenze, 2003; Kuh, Ben-Shlomo, Lynch, Hallqvist, & Power, 2003). Here, we use the Good Behavior Game (Kellam et al., 2008) prevention program as an example. A developmentally informed and universal classroom behavior management prevention model, the Good Behavior Game (Kellam et al., 2008) targets first and second grade children and has demonstrated long-term positive effects, particularly during critical periods (e.g., late adolescents and emerging adulthood), on several health outcome indicators, including substance use and delinquency. The choice of its use on entry into

elementary school was based on prospective, longitudinal studies that identified aggressive/disruptive behavior in first grade as a strong antecedent for adolescent drug and alcohol use and abuse/dependence in young adulthood (Kellam et al., 2008). Thus, the development of prevention programs such as the Good Behavior Game (Kellam et al., 2008), which aim to combat problem behaviors early and during a developmental stage that has been identified as a critical period for subsequent drug use, is an example of how developmental epidemiology may be used in informing prevention services.

Life Course/Social Field Theory

Building on developmental epidemiology, life course/social field theory posits that individuals are embedded in social fields or contexts that require social task demands. These social task demands have criteria for success or failure and vary based on developmental stages of life as well as critical transition periods across each developmental stage of life (Kellam & Van Horm, 1997). For example, learning how to drive responsibly is one social task demand that many experience during late adolescence and early adulthood in the United States. The life course/social field framework provides prevention scientists with a platform to integrate various disciplines (Kellam & Van Horm, 1997).

Genetic, psychological, environmental/ecological, developmental, and life-course factors provide a risk and protective framework for understanding complex human behaviors and the ways in which these factors both influence and are influenced by one another over time (Kaufman et al., 2007; Sloboda et al., 2012). Furthermore, descriptive epidemiology constitutes a key first step to understanding the disease or problem behavior and provides prevention scientists with the necessary tools to identify both the frequency and the distribution of risk and protective factors in populations as well as to assess the extent of a disease or problem behavior (Gordis, 2009). Analytical epidemiology then uses this information to examine, for example, these risk and protective factors to better understand the etiology of the outcome of interest. Next, we describe the role of etiological frameworks in guiding the development of preventive interventions.

The Role of Etiological Models in Prevention Science

Several etiological models have been developed to highlight the role of genetic, psychological, environmental, and developmental factors on risk behaviors, including Bronfenbrenner's (1979, 1989) ecological systems theory. From this perspective, risk behaviors are a function of both the individual and the environment. Ecological systems theory (Bronfenbrenner, 1979, 1989) conceptualizes development as taking place within contexts, namely, the microsystem, mesosystem,

exosystem, and macrosystem. The microsystem consists of risk and protective factors proximal to the individual. For example, risk factors for childhood obesity in the microsystem include the type of nutrition provided in the home. The mesosystem refers to the relationship and processes between two microsystems. Factors in the mesosystem, for example, may include parental involvement that supports healthier meal options or choices in a child's school context. The exosystem are those systems in which an individual does not directly participate but that can have an impact on the individual. For example, parental work conditions (e.g., long work hours) might prevent parents from providing a healthy dinner at home, which in turn may lead to children having to prepare processed foods and thereby have an effect on the child's eating habits. Lastly, the macrosystem encompasses those variables present in the broader social and cultural systems such as a policy requiring more transparent nutritional food labeling for consumers. Etiological models, including the ecological systems theory (Bronfenbrenner), have helped advance the field of prevention science and develop next generation etiological models.

Building on Bronfenbrenner's work, several next generation ecological frameworks have been developed to conceptualize the role of environmental/ecological factors on the etiology of risk and protection and advance the field of prevention science, including the ecodevelopmental theory (Szapocznik & Coatsworth, 1999). Ecodevelopmental theory (Szapocznik & Coatsworth, 1999), for example, is helpful in conceptualizing integrated developmental risk and protective processes operating in the lives of individuals (Pantin et al., 2003; Prado et al., 2010; Szapocznik & Coatsworth, 1999). Ecodevelopmental theory (Szapocznik & Coatsworth, 1999) affirms that social domains in which the individual is embedded both influence and are influenced by the individual in a developmental context and occurs on various levels, including those microsystems, mesosystems, exosystems, and macrosystems described above. From this viewpoint, health behaviors are influenced by a multiplicity of factors, some of which are proximal to the individual, whereas others are distal.

The development of theoretical models, such as ecodevelopmental theory (Pantin et al., 2003; Prado et al., 2010; Szapocznik & Coatsworth, 1999), serve a vital role to the field of prevention science and epidemiology. Theoretical models help prevention scientists understand etiology, aid in predicting health-risk behaviors among the populations we work with, and help ascertain the pathways through which prevention programs work (Brown et al., 2008, 2009). Conversely, the fields of prevention science and epidemiology also play an important role in developing and testing theoretical models, at both the population level and the more targeted high-risk groups.

Prevention science and epidemiology inform what population-based strategies (i.e., universal), as compared with high-risk targeted strategies (i.e., indicated), are theoretically able to accomplish (Brown & Faraone, 2004). We take, for example, the prevention of drug abuse in adolescents. LifeSkills Training, a universal school-based prevention program found to be effective in preventing/reducing drug use, was designed for all ethnic and racial students within a particular school setting

(Griffin, Botvin, Nichols, & Doyle, 2003). This drug use prevention strategy has the potential to reach a large segment of students, and therefore it is a population-based or universal strategy (Brown & Faraone, 2004). In comparison, identifying drug abuse risk factors can also lead to the development of interventions that target a specific segment of the population that shares a particular risk factor (Brown & Faraone, 2004). For example, Familias Unidas (Pantin et al., 2009; Prado et al., 2007; Prado & Pantin, 2011), a Hispanic-specific, family-based preventive intervention that targets identified risk factors that may be more pronounced in Hispanic families (e.g., cultural differences and parent–adolescent communication), has been found to be efficacious in preventing/reducing drug use among Hispanic adolescents. Thus, prevention science can be seen as an epidemiologic experiment because it aims to determine whether and the extent to which prevention programs and population-based or high-risk strategies target identified etiological and theoretical predictors, as well as identify which processes account for the differences caused by the intervention.

As shown in Fig. 1.1, epidemiology and theoretical models inform the field of prevention science and the development of preventive interventions. Prevention programs in turn can be used to test these models and help advance and adapt theoretical frameworks as a result of new evidence and knowledge gained. In the next section, we describe some of the epidemiologic methods and study designs used to evaluate prevention programs.

The Use of Epidemiologic Methods and Study Design in Evaluating Preventive Interventions

Advanced longitudinal epidemiologic methods and study designs have provided prevention scientists with the necessary tools to efficiently and more effectively evaluate preventive interventions and determine to what extent the effects are sustained over time (Hayes, 2006). These tools facilitate moving a program of research from efficacy to scale and promote the use of prevention research findings in the advancement of epidemiologic models (Kellam et al., 2011). Longitudinal methods and research designs examine change over time by following individuals beyond the period when they are actively participating in prevention or early intervention efforts and thereby provide valuable insight into the sustainability of prevention programs. Change and trajectory of outcomes over time can now be conducted because of advancements in epidemiologic data collection methods and advanced statistical methods that allow for the testing of more complex etiological and theoretical models (McArdle, 2009). Thus, advanced longitudinal methodologies inform the effects of preventive interventions, both short and long term, promote the advancement of epidemiological and prevention models, and work toward best practices. Before a prevention program has been found to be an evidence-based or best practice model of prevention, however, the program must

first undergo rigorous scientific testing and go through a research process that includes basic science, efficacy, effectiveness, implementation, and taking interventions to scale (Brown et al., 2008, 2009; Van Spall, Toren, Kiss, & Fowler, 2007). Basic science and epidemiology are used to develop etiological and theoretical models, which in turn can be used to examine the efficacy of a prevention program. The efficacy of a prevention program is determined by whether and the extent to which it works under ideal conditions. If and when a program is found to be efficacious, then effectiveness trials, which examine the effects of prevention programs in real-world settings, can take place. Once established as an effective model, prevention programs can work toward implementation and scale to ensure widespread adoption. It should not be surprising that epidemiological and prevention science methods are integrated throughout this process. For example, theoretical models may be adapted, or prevention programs tailored, based on the results of outcome and process data, which help optimize efficacious or effective models. Additionally, evaluation is particularly important throughout the research process. Several epidemiological research methods are used to evaluate the effects of prevention programs, including time to event, growth curve, multilevel modeling, mixture modeling, and mediation modeling. Below, we discuss some of these widely used methods in evaluating prevention interventions.

Of importance to the field of prevention science is the evaluation of prevention programs, and prevention trials methodology has developed as a particularly important type of experimental epidemiology (Brown et al., 2008). This is accomplished through the testing of etiological and intervention models that help guide prevention services, as well as identifying those pathways through which preventive interventions work. This is important not only to gain a better understanding of intervention processes but also to confirm that the intervention is in fact targeting the hypothesized etiological factors and is casually related to the outcomes of interest. Therefore, prevention science is concerned with, among many other things, answering the question: *How do preventive interventions work*, *for whom*, *and under what cultural*, *social*, *and institutional conditions* (Brown et al., 2008; Kellam et al., 2011; Tein, Sandler, MacKinnon, & Wolchik, 2004)?

Given that the primary goal of epidemiology is to describe the natural history of the outcome of interest, advanced statistical methods, including time-to-event, growth-curve, and multilevel models, are especially important for measurement of change over time and thereby more accurately describe the natural history process. For example, time-to-event analysis provides prevention scientists with the tools to capture, to some extent, the natural history of an outcome of interest, including the onset of the behavior of interest, progression/escalation, and the cessation/desistance of that behavior. Multilevel modeling, for example, is used to describe how individuals change over time as well as the extent to which the changes may vary across individuals (Singer & Willett, 2003). In growth-curve modeling, we can examine individual trajectories, individual differences in these trajectories, predictors of individual differences, and their effects on development over time. Furthermore, growth-curve modeling can describe important group statistics to better understand developmental processes at the group level (Duncan,

Duncan, & Strycher, 2006). For example, prevention scientists interested in examining whether and the extent to which a prevention program is efficacious in preventing an outcome of interest, relative to a control condition, may find the growth-curve modeling approach useful. Here, we can describe trajectories of individuals both within and across conditions. Additionally, mixture modeling, including latent class growth analysis (LCGA) and growth mixture modeling (GMM), allows prevention scientists to capture classes or subpopulations that are unknown (Muthen & Muthen, 1998–2010). For example, mixture modeling may be helpful in identifying those classes or subpopulations that benefit most from participating in a drug abuse prevention program.

Some research has been conducted to highlight for whom preventive interventions work and for whom they do not work. For example, Hispanic adolescent preventive interventions may be more efficacious among US-born youth, relative to foreign-born youth, on some behavioral and mental health outcomes (i.e., moderator; Cordova et al., 2011; Martinez & Eddy, 2005). Additionally, research has demonstrated the ways in which the effects of preventive interventions vary by gender (i.e., moderator; Kulis, Marsiglia, Ayers, Calderón-Tena, & Nuño-Gutiérrez, 2011; Kulis, Yabiku, Marsiglia, Nieri, & Crossman, 2007). Furthermore, research has been conducted to identify how prevention programs work. For example, we now know that family functioning is one pathway through which family-based, drug abuse preventive interventions are efficacious (i.e., mediator; Pantin et al., 2009; Prado et al., 2007; Sandler, Schoenfelder, Wolchik, & MacKinnon, 2011). Although some research has focused on answering for whom and how preventive interventions work, little research has been done to identify research processes associated with the cultural, social, and institutional conditions that may have an effect on prevention programs (Kellam et al., 2011; Sandler et al., 2011). Working toward a better understanding of these research processes might have great utility in identifying potential factors that contribute to prevention intervention outcomes. This knowledge may elucidate optimally efficacious and effective preventive interventions, which in turn could aid in adapting or extending etiological and theoretical models (Kellam et al., 2011). One widely used model for better understanding research processes and testing theoretical frameworks in the field of prevention science is the mediation model.

The mediation model examines whether and the extent to which a third variable has an effect on the relation between two other variables (Brug, Oenema, & Ferreira, 2005). Theoretical frameworks are especially important in the selection of mediating processes and predictors because they, by definition, are informed by a large body of scientific research and can provide strong evidence for the etiology (e.g., risk and protective factors) of a disease. In addition to facilitating the selection of mediation processes, theoretical frameworks present a systematic way of understanding causal determinants of disease and how these factors may operate to reduce disease, which may be helpful in advancing the fields of epidemiology and prevention science. Once a viable mediation process has been identified, a hypothesis can then be generated that will guide the research protocol, including the pathways (e.g., family functioning) to target and how change processes occur.

Thus, a mediation analysis can test a theory in intervention; to the extent that changes in a hypothesized mediator predict distal outcomes, the theoretical model is confirmed (MacKinnon & Luecken, 2008).

In addition to the importance of statistical methods in evaluating preventive interventions, equally important is the study design. Below, we describe several study designs that are helpful in evaluating prevention programs.

The randomized controlled trial (RCT) can be particularly helpful in moving forward the prevention science field because RCTs, in theory, generate equivalent condition groups that in turn allow prevention scientists to say with greater certainty that prevention effects are attributable to the intervention itself (Jadad et al., 1996; Olsen, Christensen, Murray, & Ekbom, 2010; Silverman, 2009). Although RCTs are the optimal strategy in the field of prevention science, they may not be feasible in the evaluation of all preventive interventions (Silverman, 2009). For example, there may be circumstances in which participants cannot plausibly and ethically be randomized to an experimental or control condition; consider faith-based interventions among persons who may not practice organized religion (Rubin, 1974, 2005). Therefore, several preventive intervention evaluation alternatives exist that may be more appropriate for examining the effects of prevention programs, including observational, rolled-out, quasi-experimental, and randomized encouragement designs.

Observational methods may be particularly helpful for better understanding disease in instances where randomization of participants is impossible or unethical or in the outbreak of a rare disease. Observational methods, including the cohort design, allow prevention scientists to follow a segment of the population that may not have the disease or disorder of interest and observe this sample as a natural ecological experiment. That is, through longitudinal and correlation analyses and life histories, the cohort design allows prevention scientists to examine risk and protective factors associated with that particular disease or disorder in a natural setting (Hemingway & Marmot, 1999; Stroup et al., 2000). A limitation to observational studies, however, is that they introduce several types of biases, including selection bias when the choice of control groups may differ systematically from intervention groups (Benson & Hartz, 2000; Hemingway & Marmot, 1999).

Another alternative to the traditional RCT design is the rolled-out or dynamic wait list design (Brown, Wyman, Guo, & Peña, 2006; Brown et al., 2009). In this design, an intervention that has already been selected for wide-scale implementation is evaluated. For instance, a government mandate may require that all schools within a specific district deliver a sex education program. The sex education program may have limited empirical support, but an opportunity exists to evaluate the program as it is "rolled-out" into the school district. In the rolled-out methodology, all of the study units (individuals or clusters) receive the intervention but are randomly assigned to a time interval in which to receive the intervention. More specifically, all of the study units begin in the control condition and, as participants receive the preventive intervention services, participants are switched to the experimental condition. This process occurs until all of the study units eventually receive the intervention. Finally, the effects of the experimental condition on the outcomes

are examined and compared with the control condition. This process occurs until the last interval, when all study units have received the intervention and comparisons can no longer be made. There are several benefits to this type of design. For instance, the rolled-out method is advantageous in circumstances where ethical concerns or public health mandates do not support the exposure of an intervention to only a subset of individuals. The rolled-out method is also advantageous for practical reasons, for instance, when financial resources limit the ability to deliver the intervention simultaneously to all study participants (Brown et al., 2006, 2009).

Nonrandom or quasi-experimental trials are also alternatives to RCTs. Several types of quasi-experimental designs exist. For example, in time-series quasi-designs, a single group of study participants is assessed both before and after the prevention intervention activities. A significant favorable outcome between pre- and post-intervention, to the extent that it is not due to chance, suggests an efficacious/effective intervention (Grimshaw, Campbell, Eccles, & Steen, 2000).

Randomized encouragements are epidemiologic designs in which participants are randomized not to a preventive intervention but rather to an opportunity or encouragement to take part in an intervention (Vinokur, Price, & Schul, 1995). In a smoking cessation prevention program, for example, participants may be randomly assigned either to receive an invitation or not to receive an invitation. In this methodology, it is assumed that the impact of the assignment is mediated entirely through the receipt of the treatment. This process aims to remove the necessity to fully adhere to the prevention protocol, as is the case in RCTs. Finally, in nonrandom quantitative assignment designs, participants are not randomized but rather assigned to treatment based on a quantitative measure such as financial need (Finkelstein, Levin, & Robbins, 1996a, 1996b).

Additionally, epidemiological study designs include case control studies, cohort-age-period designs, twin studies, and cross designs. Case control studies, for example, have great utility in situations where the disease of interest is rare, observing a sample or "control series" of the population. This sample is then used in place of the complete assessment of disease frequencies, including the proportion, rate, and odds (Rothman, Greenland, & Rash, 2008).

Applying Prevention Research Findings to Epidemiology: Informing the Advances of Epidemiologic Models

We have described the ways in which epidemiological data have great utility guiding and informing the development and evaluation of preventive interventions. For example, epidemiological data show that HIV risk behaviors are a complex phenomena occurring at multiple levels (DiClemente et al., 2004; Prado, Pantin, Schwartz, Lupei, & Szapocznik, 2006), which in turn has called for the development and evaluation of HIV preventive interventions at multiple levels, including family (Villarruel, Jemmott, Jemmott, & Eakin, 2006), peer (Guarino, Deren, Mino,

Sung-Yeon, & Shedlin, 2010; Raja, McKirnan, & Glick, 2007), school (Nkansah-Amankra, Diedhiou, Agbanu, Harrod, & Dhawan, 2011), and community (Harper, Bangi, Sanchez, Doll, & Pedraza, 2009) to help curb the tide of HIV infection. Although the use of epidemiologic findings to inform the development and evaluation of preventive interventions has been widely practiced, little has been done with respect to using prevention findings to inform epidemiology (Sloboda et al., 2012). That is, the use of scientific knowledge produced from the implementation and evaluation of prevention programs to inform epidemiology has been minimal. Working toward epidemiologic models that are informed by prevention findings might prove useful in advancing the fields of epidemiology and prevention science, particularly with respect to public health. Therefore, the contribution of prevention science for epidemiology is then to better specify when to measure, what to measure, whom to measure, what to target, and more specifically to assess the viability of the dominant perspective on the etiology of the natural history of the outcome.

To highlight one way that prevention findings can inform epidemiology, we refer back to the mediation model and use the Familias Unidas (Pantin et al., 2009; Prado et al., 2007; Prado & Pantin, 2011) preventive program as an example. As previously mentioned, family functioning is the pathway through which Familias Unidas has an effect on both drug use and unprotected sex in multiple efficacy studies (Pantin et al., 2009; Prado et al., 2007; Prado & Pantin, 2011). The question still remains, however: *Will family functioning remain the significant pathway through which Familias Unidas operates if/when the program moves to implementation*? For example, it may very well be that other factors such as community processes, including community partnerships and networks, may be the pathway through which the Familias Unidas program has an effect on the outcomes. This in turn can be used to inform epidemiologic models. To that end, Kellam and colleagues (2011) have proposed several strategies that may be helpful to the process of prevention findings informing epidemiologic models and the field of prevention science, including community partnerships and networks (Kellam et al., 2011).

Community partnerships are essential to the research process and for the development of the next generation of prevention research. For example, preventive interventions are effective only to the extent that they are accepted and implemented with fidelity by the community. Thus, building and sustaining community partnerships for prevention efforts are strategies through which prevention science findings may inform epidemiologic models, which in turn may advance the field of prevention science (Kellam et al., 2011). For example, community partnerships can inform prevention scientists whether and the extent to which the prevention program's aims and goals are shared with that of the communities' aims and goals (Kellam et al., 2011). This information can then be used to tailor prevention services and epidemiological models and work toward optimally efficacious and effective programs.

Prevention research findings may also help advance epidemiologic models, particularly when moving a prevention program from efficacy to effectiveness to

implementation (Kellam et al., 2011). To do so, networks should be established that consist of researchers, policy makers, and practitioners with a shared goal of working toward translational research (i.e., efficacy to effectiveness to implementation). To this end, networks can foster the sharing of diverse research experiences, including identified theoretical models, measures, and relevant assessment for a particular population and context (Kellam et al., 2011).

A prevention model that has been shown to be effective, particularly as it relates to implementation strategies, is the RE-AIM framework (Glasgow, McKay, Piette, & Reynolds, 2001; Glasgow, Vogt, & Boles, 1999). The RE-AIM framework (Glasgow et al., 1999, 2001) is a translational research model—moving efficacious programs to real-world settings—and is concerned with factors and processes related to long-term effectiveness impact. The RE-AIM framework is guided by five steps: reach, efficacy, adoption, implementation, and maintenance (Glasgow et al., 1999, 2001). Reach refers to engagement, active participation, and retention in prevention services, which are affected by contextual barriers to participation, including transportation, work schedule, and child care. Efficacy is concerned with the effects of a prevention program on the outcome variables when it is implemented with fidelity and in a controlled setting. Adoption can be described as the process by which a best practice program is delivered at the system level and the proportion of organizations that are willing to adopt a particular best practice model. Implementation operates at the real-world level and is concerned with the consistency of delivery and translation of a best practice model. Finally, maintenance is concerned with the long-term effects of prevention programs on behavior and how the best practice is institutionalized and put into practice (Glasgow et al., 1999, 2001). The Nurse Family Partnership (Olds, 2008; Olds, Henderson, & Kitzman, 1994; Olds, Henderson, Kitzman, & Cole, 1995) prevention program is an example of a translational program of research that has implemented the RE-AIM framework and has been rigorously tested over the past 20 years (Donelan-McCall, Eckenrode, & Olds, 2009). Nurse Family Partnership (Olds, 2008; Olds et al., 1994, 1995), a home-visiting prevention intervention, is efficacious in preventing/reducing child maltreatment, including child abuse and neglect, emergency room visits, and mothers' substance use during gestation. First evaluated in 1977 in an RCT, Nurse Family Partnership has spanned a program of research from reach to efficacy to adoption to implementation to maintenance (Donelan-McCall et al., 2009). In fact, the Nurse Family Partnership now houses a national service office that assists communities with replication and ensuring fidelity in the implementation of the Nurse Family Partnership program. Here, the Nurse Family Partnership provides a model to the field of prevention science on how prevention findings can be used to inform epidemiology as it moves along the research continuum from basic science to implementation.

Qualitative methodologies may play an important role in informing the feedback loop between prevention science and epidemiology. Although the use of qualitative methods in the fields of prevention science and epidemiology has been minimal, prevention scientists have developed an increased awareness with respect to the use of qualitative methods, particularly to understanding processes at the

implementation stage (Bucher Della Torre, Akré, & Suris, 2010; Pontin, Peters, Lobban, Rogers, & Morriss, 2009; Voogdt-Pruis, Beusmans, Gorgels, & van Ree, 2011). To this end, qualitative methods have great utility in elucidating processes that facilitate or hinder adoption of effective programs, including feasibility and acceptability of prevention strategies among individuals and communities (Parra-Cardona, Cordova, Holtrop, Villarruel, & Wieling, 2008; Parra Cardona et al., 2009), contextual challenges experienced by underserved populations that could serve as barriers to participation (Cervantes & Cordova, 2011; Cordova & Cervantes, 2010), and adaptations to ensure adoption of prevention services in communities (Dodson et al., 2009). Thus, qualitative methodology, when used as a tool to gather prevention findings and process data, has the potential to serve as a feedback loop and inform epidemiology.

To summarize, the field of prevention science could benefit from using prevention findings in the advancement of epidemiologic models. Doing so may prove helpful in transporting a prevention program from bench to practice. As shown in Fig. 1.1, the prevention research cycle should include a feedback loop consisting of multiple steps, including the identification of a problem and large-scale implementation.

Epidemiology and Prevention Science: Working Toward a Mutually Informed Process

In summary, epidemiology plays a significant role in the field of prevention science. Central to epidemiology is the description of the natural history or career of the outcome of interest and consequently focuses on factors that contribute to the onset, progression/escalation, and cessation/desistance of the outcome of interest. Epidemiology provides a framework for prevention scientists to identify the outcome of interest and develop etiologic models and preventive interventions, methods, and study designs to test preventive interventions. Such efforts are aimed at promoting health and well-being in individuals, families, and communities alike. Although a substantial amount of research has demonstrated the ways in which epidemiological findings are useful to the field of prevention science, little has been done with regard to prevention findings informing epidemiologic models. Working toward a mutually informed process will advance the fields of epidemiology and prevention science. A mutually informed process, whereby both epidemiology and prevention science inform each other, may aid in the development of optimally efficacious and effective preventive interventions that promote health and prevent/reduce health-risk behaviors.

Acknowledgement The preparation of this chapter was supported by National Institute on Drug Abuse (NIDA) grant # R01DA025192-02S1, NIDA and National Institute on Alcohol Abuse and Alcoholism grant #3R01DA025192-01A1S1 awarded to Guillermo Prado, and National Institute on Minority Health and Health Disparities grant #1L60MD006269-01 awarded to David Cordova.

References

Behrendt, S., Wittchen, H. U., Höfler, M., Lieb, R., & Beesdo, K. (2009). Transitions from first substance use to substance use disorders in adolescence: Is early onset associated with a rapid escalation? *Drug and Alcohol Dependence, 99*(1–3), 68–78.

Benson, K., & Hartz, A. J. (2000). A comparison of observational studies and randomized, controlled trials. *New England Journal of Medicine, 342*(25), 1878–1886. doi:10.1056/NEJM200006223422506.

Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, W. S., & Weikart, D. P. (1987). The effects of early educational intervention on crime and delinquency in adolescence and early adulthood. In J. D. Burchard & S. N. Burchard (Eds.), *Prevention of delinquent behavior* (Vol. X, pp. 220–240). Newbury Park, CA: Sage Publications.

Blackburn, A. C., & Jerry, D. J. (2011). Map making in the 21st century: Charting breast cancer susceptibility pathways in rodent models. *Journal of Mammary Gland Biology and Neoplasia, 16*, 57–64.

Botvin, G. J., Griffin, K. W., Paul, E., & Macaulay, A. P. (2003). Preventing tobacco and alcohol use among elementary school students through life skills training. *Journal of Child and Adolescent Substance Abuse, 12*(4), 1–7. doi:10.1300/J029v12n04_01.

Braveman, P., & Barclay, C. (2009). Health disparities beginning in childhood: A life-course perspective. *Pediatrics, 124*(Suppl 3), S163–S175. doi:10.1542/peds.2009-1100D.

Brody, G. H., Beach, S. R., Philibert, R. A., Chen, Y. F., & Murry, V. M. (2009). Prevention effects moderate the association of 5-HTTLPR and youth risk behavior initiation: Gene x environment hypotheses tested via a randomized prevention design. *Child Development, 80*(3), 645–661. doi:10.1111/j.1467-8624.2009.01288.x.

Brody, G. H., Chen, Y. F., Beach, S. R., Philibert, R. A., & Kogan, S. M. (2009). Participation in a family-centered prevention program decreases genetic risk for adolescents' risky behaviors. *Pediatrics, 124*(3), 911–917. doi:10.1542/peds.2008-3464.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (1989). Ecological systems theory. In R. Vasta (Ed.), *Annals of child development* (Vol. 6, pp. 18–249). Greenwich, CT: JAI Press.

Brown, C. H., & Faraone, S. V. (2004). Prevention of schizophrenia and psychotic behavior: Definitions and methodologic issues. In W. S. Stone, S. V. Faraone, & M. T. Tsuang (Eds.), *Early clinical intervention and prevention in schizophrenia* (pp. 255–284). Totowa, NJ: The Humana Press.

Brown, C. H., Ten Have, T. R., Jo, B., Dagne, G., Wyman, P. A., Muthén, B., et al. (2009). Adaptive designs for randomized trials in public health. *Annual Review of Public Health, 30*(1), 1–25. doi:10.1146/annurev.publhealth.031308.100223.

Brown, C. H., Wang, W., Kellam, S. G., Muthén, B. O., Petras, H., Toyinbo, P., et al. (2008). Methods for testing theory and evaluating impact in randomized field trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug and Alcohol Dependence, 95*(Suppl 1), S74–S104. doi:10.1016/j.drugalcdep.2007.11.013.

Brown, C. H., Wyman, P. A., Guo, J., & Peña, J. (2006). Dynamic wait-listed designs for randomized trials: New designs for prevention of youth suicide. *Clinical Trials, 3*(3), 259–271. doi:10.1191/1740774506cn152oa.

Brug, J., Oenema, A., & Ferreira, I. (2005). Theory, evidence and intervention mapping to improve behavior nutrition and physical activity interventions. *International Journal of Behavioral Nutrition and Physical Activity, 2*(1), 2.

Bucher Della Torre, S., Akré, C., & Suris, J. C. (2010). Obesity prevention opinions of school stakeholders: A qualitative study. *Journal of School Health, 80*(5), 233–239. doi:10.1111/j.1746-1561.2010.00495.x.

Catalano, R. F., & Hawkins, J. D. (1996). The social development model: A theory of antisocial behavior. In J. D. Hawkins (Ed.), *Delinquency and crime: Current theories* (pp. 149–197). New York, NY: Cambridge University Press.

Cervantes, R. C., & Cordova, D. (2011). Life experiences of Hispanic adolescents: Developmental and language considerations in acculturation stress. *Journal of Community Psychology, 39*(3), 336–352. doi:10.1002/jcop.20436.

Chambers, R. A., Taylor, J. R., & Potenze, M. N. (2003). Developmental neurocircuitry of motivation in adolescence: A critical period of addiction vulnerability. *American Journal of Psychiatry, 160*(6), 1041–1052.

Cleveland, M. J., Feinberg, M. E., & Greenberg, M. T. (2010). Protective families in high- and low-risk environments: Implications for adolescent substance use. *Journal of Youth and Adolescence, 39*(2), 114–126. doi:10.1007/s10964-009-9395-y.

Cordova, D., & Cervantes, R. C. (2010). Intergroup and within-group perceived discrimination among U.S.-born and foreign-born Latino youth. *Hispanic Journal of Behavior Science, 32*(2), 259–274. doi:10.1177/0739986310362371.

Cordova, D., Huang, S., Arzon, M., Freitas, D., Malcolm, S., & Prado, G. (2011). The role of attitudes, family, peer and school on alcohol use, rule breaking and aggressive behavior in Hispanic delinquent adolescents. *Open Family Studies Journal, 4*, 38–45.

DiClemente, R. J., Wingood, G. M., Harrington, K. F., Lang, D. L., Davies, S. L., Hook, E. W., 3rd, et al. (2004). Efficacy of an HIV prevention intervention for African American adolescent girls. *JAMA, 292*(2), 171–179. doi:10.1001/jama.292.2.171.

Dishion, T. J., Kavanagh, K., Schneiger, A., Nelson, S., & Kaufman, N. K. (2002). Preventing early adolescent substance use: A family-centered strategy for the public middle school. *Prevention Science, 3*(3), 191–201.

Dodson, E. A., Fleming, C., Boehmer, T. K., Haire-Joshu, D., Luke, D. A., & Brownson, R. C. (2009). Preventing childhood obesity through state policy: Qualitative assessment of enablers and barriers. *Journal of Public Health Policy, 30*(S1), S161–S176.

Donelan-McCall, N., Eckenrode, J., & Olds, D. L. (2009). Home visiting for the prevention of child maltreatment: Lessons learned during the past 20 years. *Pediatric Clinics of North America, 56*(2), 389–403. doi:10.1016/j.pcl.2009.01.002.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum Associates. Revised Edition.

Finkelstein, M. O., Levin, B., & Robbins, H. (1996a). Clinical and prophylactic trials with assured new treatment for those at greater risk: I, A design proposal. *American Journal of Public Health, 86*(5), 691–695.

Finkelstein, M. O., Levin, B., & Robbins, H. (1996b). Clinical and prophylactic trials with assured new treatment for those at greater risk: II, Examples. *American Journal of Public Health, 86*(5), 696–705. doi:10.2105/ajph.86.5.696.

Franks, P. W., & Poveda, A. (2011). Gene-lifestyle and gene-pharmacotherapy interactions in obesity and its cardiovascular consequences. *Current Vascular Pharmacology, 9*(4), 401–456.

Glasgow, R. E., McKay, H. G., Piette, J. D., & Reynolds, K. D. (2001). The RE-AIM framework for evaluating interventions: What can it tell us about approaches to chronic illness management? *Patient Education and Counseling, 44*(2), 119–127. doi:10.1016/s0738-3991(00)00186-5.

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health, 89*(9), 1322–1327. doi:10.2105/ajph.89.9.1322.

Gordis, L. (2009). *Epidemiology* (4th ed.). Philadelphia, PA: Saunders.

Grant, B. F., & Dawson, D. A. (1998). Age at 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*, 163–173.

Griffin, K. W., & Botvin, G. J. (2010). Evidence-based interventions for preventing substance use disorders in adolescents. *Child and Adolescent Psychiatric Clinics of North America, 19*(3), 505–526. doi:10.1016/j.chc.2010.03.005.

Griffin, K. W., Botvin, G. J., Nichols, T. R., & Doyle, M. M. (2003). Effectiveness of a universal drug abuse prevention approach for youth at high risk for substance use initiation. *Preventive Medicine, 36*(1), 1–7.

Grimshaw, J., Campbell, M., Eccles, M., & Steen, N. (2000). Experimental and quasi-experimental designs for evaluating guideline implementation strategies. *Family Practice, 17* (Suppl 1), S11–S16.

Guarino, H., Deren, S., Mino, M., Sung-Yeon, K., & Shedlin, M. G. (2010). Training drug treatment patients to conduct peer-based HIV outreach: An ethnographic perspective on peers' experiences. *Substance Use & Misuse, 45*(3), 414–436.

Harper, G. W., Bangi, A. K., Sanchez, B., Doll, M., & Pedraza, A. (2009). A quasi-experimental evaluation of a community-based HIV prevention intervention for Mexican American female adolescents: The SHERO's program. *AIDS Education and Prevention, 21*, 109–123.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112*(1), 64–105. doi:10.1037/0033-2909.112.1.64.

Hawkins, J. D., Oesterle, S., Brown, E. C., Monahan, K. C., Abbott, R. D., Arthur, M. W., et al. (2012). Sustained decreases in risk exposure and youth problem behaviors after installation of the Communities That Care prevention system in a randomized trial. *Archives of Pediatrics & Adolescent Medicine, 166*(2), 141–148. doi:10.1001/archpediatrics.2011.183.

Hayes, A. (2006). Maintaining the gains sustainability in prevention and early intervention. *Family Matters, 75*, 66–69.

Heath, A. C., Bucholz, K. K., Madden, P. A., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., et al. (1997). Genetic and environmental contributions to alcohol dependence risk in a national twin sample: Consistency of findings in women and men. *Psychological Medicine, 27*(6), 1381–1396.

Helbig, K. L., Nothnagel, M., Hampe, J., Balschun, T., Nikolaus, S., Schreiber, S., et al. (2012). A case-only study of gene-environment interaction between genetic susceptibility variants in *NOD2* and cigarette smoking in Crohn's disease aetiology. *BMC Medical Genetics, 113*, 14. doi:10.1186/1471-2350-13-14.

Hemingway, H., & Marmot, M. (1999). Psychosocial factors in the aetiology and prognosis of coronary heart disease: Systematic review of prospective cohort studies. *BMJ, 318*(7196), 1460–1467. doi:10.1136/bmj.318.7196.1460.

Hemphill, S. A., Heerde, J. A., Herrenkohl, T. I., Patton, G. C., Toumbourou, J. W., & Catalano, R. F. (2011). Risk and protective factors for adolescent substance use in Washington state, the United States and Victoria, Australia: A longitudinal study. *Journal of Adolescent Health, 49* (3), 312–320. doi:10.1016/j.jadohealth.2010.12.017.

Hops, H., Ozechowski, T. J., Waldron, H. B., Davis, B., Turner, C. W., Brody, J. L., et al. (2011). Adolescent health-risk sexual behaviors: Effects of a drug abuse intervention. *AIDS and Behavior, 15*(8), 1664–1676.

Howe, G. W., Beach, S. R., & Brody, G. H. (2010). Microtrial methods for translating gene-environment dynamics into preventive interventions. *Prevention Science, 11*(4), 343–354.

Ibrahim, J. K., Anderson, E. D., Burris, S. C., & Wagenaar, A. C. (2011). State laws restricting driver use of mobile communications devices: Distracted-driving provisions, 1992–2010. *American Journal of Preventive Medicine, 40*(6), 659–665. doi:10.1016/j.amepre.2011.02.024.

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., et al. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control Clinical Trials, 17*(1), 1–12.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. (2013). *Monitoring the future national results on drug use: 2012 Overview, key findings on adolescent drug use*. Ann Arbor, MI: Institute for Social Research, The University of Michigan.

Kaufman, J., Yang, B. Z., Douglas-Palumberi, H., Crouse-Artus, M., Lipschitz, D., Krystal, J. H., et al. (2007). Genetic and environmental predictors of early alcohol use. *Biological Psychiatry, 61*(11), 1228–1234. doi:10.1016/j.biopsych.2006.06.039.

Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., et al. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95* (Suppl 1), S5–S28. doi:10.1016/j.drugalcdep.2008.01.004.

Kellam, S. G., & Langevin, D. J. (2003). A framework for understanding "evidence" in prevention research and programs. *Prevention Science, 4*(3), 137–153. doi:10.1023/a:1024693321963.

Kellam, S. G., Mackenzie, A. C., Brown, C. H., Poduska, J. M., Wang, W., Petras, H., et al. (2011). The good behavior game and the future of prevention and treatment. *Addiction Science & Clinical Practice, 6*(1), 73–84.

Kellam, S. G., & Van Horm, Y. V. (1997). Life course development, community epidemiology, and preventive trials: A scientific structure for prevention research. *American Journal of Community Psychology, 25*(4), 177–188.

Kellam, S. G., Werthamer-Larsson, L., Dolan, L. J., Brown, C. H., Mayer, L. S., Rebok, G. W., et al. (1991). Developmental epidemiologically based preventive trials: Baseline modeling of early target behaviors and depressive symptoms. *American Journal of Community Psychology, 19*(4), 563–584.

Kuh, D., Ben-Shlomo, Y., Lynch, J., Hallqvist, J., & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health, 57*(10), 778–783.

Kulis, S., Marsiglia, F., Ayers, S., Calderón-Tena, C., & Nuño-Gutiérrez, B. (2011). Gender differences in drug resistance skills of youth in Guanajuato, Mexico. *Journal of Primary Prevention, 32*(2), 113–127. doi:10.1007/s10935-011-0239-7.

Kulis, S., Yabiku, S. T., Marsiglia, F. F., Nieri, T., & Crossman, A. (2007). Differences by gender, ethnicity, and acculturation in the efficacy of the Keepin' it REAL model prevention program. *Journal of Drug Education, 37*(2), 123–144.

MacKinnon, D. P., & Luecken, L. J. (2008). How and for whom? Mediation and moderation in health psychology. *Health Psychology, 27*(2), S99–S100. doi:10.1037/0278-6133.27.2 (Suppl.).S99.

Martinez, C. R., Jr., & Eddy, J. M. (2005). Effects of culturally adapted parent management training on Latino youth behavioral health outcomes. *Journal of Consulting and Clinical Psychology, 73*(5), 841–851.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*(1), 577–605. doi:10.1146/annurev.psych.60.110707.163612.

Muthen, L. K., & Muthen, B. O. (1998–2010). *Mplus user's guide*. Los Angeles, CA: Author.

National Highway Traffic Safety Administration. (2009). An examination of driver distraction as recorded in NHTSA databases.

Nkansah-Amankra, S., Diedhiou, A., Agbanu, H., Harrod, C., & Dhawan, A. (2011). Correlates of sexual risk behaviors among high school students in Colorado: Analysis and implications for school-based HIV/AIDS programs. *Maternal and Child Health Journal, 15*(6), 730–741. doi:10.1007/s10995-010-0634-3.

Olds, D. L. (2008). Preventing child maltreatment and crime with prenatal and infancy support of parents: The nurse-family partnership. *Journal of Scandinavian Studies in Criminology and Crime Prevention, 9*(Suppl 1), 2–24. doi:10.1080/14043850802450096.

Olds, D. L., Henderson, C. R., & Kitzman, H. (1994). Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25 to 50 months of life? *Pediatrics, 93*(1), 89–98.

Olds, D. L., Henderson, C. R., Kitzman, H., & Cole, R. (1995). Effects of prenatal and infancy nurse home visitation on surveillance of child maltreatment. *Pediatrics, 95*(3), 365–372.

Olsen, J., Christensen, K., Murray, J., & Ekbom, A. (2010). *The randomized controlled trial (RCT): An introduction to epidemiology for health professionals* (Vol. 1, pp. 81–84). New York, NY: Springer.

Pantin, H., Coatsworth, J. D., Feaster, D. J., Newman, F. L., Briones, E., Prado, G., et al. (2003). Familias Unidas: The efficacy of an intervention to promote parental investment in Hispanic immigrant families. *Prevention Science, 4*(3), 189–201.

Pantin, H., Prado, G., Lopez, B., Huang, S., Tapia, M. I., Schwartz, S. J., et al. (2009). A randomized controlled trial of Familias Unidas for Hispanic adolescents with behavior problems. *Psychosomatic Medicine, 71*(9), 987–995. doi:10.1097/PSY.0b013e3181bb2913.

Parra Cardona, J., Holtrop, K., Córdova, J. D., Escobar-Chew, A. R., Horsford, S., Tams, L., et al. (2009). "Queremos Aprender": Latino immigrants' call to integrate cultural adaptation with best practice knowledge in a parenting intervention. *Family Process, 48*(2), 211–231. doi:10.1111/j.1545-5300.2009.01278.x.

Parra-Cardona, J. R., Cordova, D., Holtrop, K., Villarruel, F. A., & Wieling, E. (2008). Shared ancestry, evolving stories: Similar and contrasting life experiences described by foreign born and U.S. born Latino parents. *Family Process, 47*(2), 157–172. doi:10.1111/j.1545-5300.2008.00246.x.

Pontin, E., Peters, S., Lobban, F., Rogers, A., & Morriss, R. (2009). Enhanced relapse prevention for bipolar disorder: A qualitative investigation of value perceived for service users and care coordinators. *Implementation Science, 4*(1), 4.

Prado, G. J., Huang, S., Maldonado-Molina, M., Bandiera, F., Schwartz, S. J., de la Vega, P., et al. (2010). An empirical test of ecodevelopmental theory in predicting HIV risk behaviors among Hispanic youth. *Health Education & Behavior, 37*(1), 97–114. doi:10.1177/1090198109349218.

Prado, G. J., & Pantin, H. (2011). Reducing substance use and HIV health disparities among Hispanic youth: The Familias Unidas program of research. *Psychosocial Intervention, 20*(1), 63–73.

Prado, G. J., Pantin, H., Briones, E., Schwartz, S. J., Feaster, D., Huang, S., et al. (2007). A randomized controlled trial of a parent-centered intervention in preventing substance use and HIV risk behaviors in Hispanic adolescents. *Journal of Consulting and Clinical Psychology, 75*(6), 914–926. doi:10.1037/0022-006x.75.6.914.

Prado, G. J., Pantin, H., Huang, S., Cordova, D., Tapia, M. I., Velazquez, M.-R., et al. (2012). Effects of a family intervention in reducing HIV risk behaviors among high-risk Hispanic adolescents: A randomized controlled trial. *Archives of Pediatrics & Adolescent Medicine, 166*(2), 127–133. doi:10.1001/archpediatrics.2011.189.

Prado, G. J., Pantin, H., Schwartz, S. J., Lupei, N. S., & Szapocznik, J. (2006). Predictors of engagement and retention into a parent-centered, ecodevelopmental HIV preventive intervention for Hispanic adolescents and their families. *Journal of Pediatric Psychology, 31*(9), 874–890. doi:10.1093/jpepsy/jsj046.

Prado, G. J., Schwartz, S. J., Maldonado-Molina, M., Huang, S., Pantin, H. M., Lopez, B., et al. (2009). Ecodevelopmental x intrapersonal risk: Substance use and sexual behavior in Hispanic adolescents. *Health Education & Behavior, 36*(1), 45–61. doi:10.1177/1090198107311278.

Raja, S., McKirnan, D., & Glick, N. (2007). The treatment advocacy program - Sinai: A peer-based HIV prevention intervention for working with African American HIV-infected persons. *AIDS and Behavior, 11*, 127–137. doi:10.1007/s10461-007-9226-7.

Rothman, K. J., Greenland, S., & Rash, T. L. (2008). *Modern epidemiology*. Philadelphia, PA: Lippincott Williams Wilkins.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*(5), 688–701. doi:10.1037/h0037350.

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association, 100*(469), 322–331.

Sandler, I. N., Schoenfelder, E. N., Wolchik, S. A., & MacKinnon, D. P. (2011). Long-term impact of prevention programs to promote effective parenting: Lasting effects but uncertain processes. *Annual Review of Psychology, 62*(1), 299–329. doi:10.1146/annurev.psych.121208.131619.

Schwartz, S. J., Pantin, H., Coatsworth, J. D., & Szapocznik, J. (2007). Addressing the challenges and opportunities for today's youth: Toward an integrative model and its implications for research and intervention. *Journal of Primary Prevention, 28*(2), 117–144. doi:10.1007/s10935-007-0084-x.

Silverman, S. (2009). From randomized controlled trials to observational studies. *American Journal of Medicine, 122*(2), 114–120. doi:10.1016/j.amjmed.2008.09.030.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Sloboda, Z., Glantz, M., & Tarter, R. E. (2012). Revisiting the concepts of risk and protective factors for understanding the etiology and development of substance use and substance use disorders: Implications for prevention. *Substance Use and Misuse, 47*, 944–962.

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., et al. (2000). Meta-analysis of observational studies in epidemiology. *JAMA, 283*(15), 2008–2012. doi:10.1001/jama.283.15.2008.

Susser, M., & Stein, Z. (2009). *Eras in epidemiology: The evolution of ideas*. New York, NY: Oxford University Press.

Swendsen, J., Burstein, M., Case, B., Conway, K. P., Dierker, L., He, J., et al. (2012). Use and abuse of alcohol and illicit drugs in U.S. adolescents: Results of the National Comorbidity Survey – Adolescent supplement. *Archive of General Psychiatry, 69*(4), 390–398.

Szapocznik, J., & Coatsworth, J. D. (1999). An ecodevelopmental framework for organizing the influences on drug abuse: A developmental model of risk and protection. In M. Glantz & C. R. Hartel (Eds.), *Drug abuse: Origins and interventions* (pp. 331–366). Washington, DC: American Psychological Association.

Tein, J.-Y., Sandler, I. N., MacKinnon, D. P., & Wolchik, S. A. (2004). How did it work? Who did it work for? Mediation in the context of a moderated prevention effect for children of divorce. *Journal of Consulting and Clinical Psychology, 72*(4), 617–624. doi:10.1037/0022-006x.72.4.617.

Tonin, S. L., Burrow-Sanchez, J. J., Harrison, R. S., & Kircher, J. C. (2008). The influence of attitudes, acculturation, and gender on substance use for Mexican American middle school students. *Addictive Behaviors, 33*(7), 949–954. doi:10.1016/j.addbeh.2008.02.014.

Torrence, M. E. (1997). *Mosby's biomedical science series: Understanding epidemiology*. St. Louis, MO: Academic Press.

Van Spall, H. G. C., Toren, A., Kiss, A., & Fowler, R. A. (2007). Eligibility criteria of randomized controlled trials published in high-impact general medical journals. *JAMA, 297*(11), 1233–1240. doi:10.1001/jama.297.11.1233.

Vega, W. A., Ang, A., Rodriguez, M. A., & Finch, B. K. (2011). Neighborhood protective effects on depression in Latinos. *American Journal of Community Psychology, 47*(12), 114–126. doi:10.1007/s10464-010-9370-5.

Villarruel, A. M., Jemmott, L. S., Jemmott, J. B., & Eakin, B. L. (2006). Recruitment and retention of Latino adolescents to a research study: Lessons learned from a randomized clinical trial. *Journal for Specialists in Pediatric Nursing, 11*(4), 244–250. doi:10.1111/j.1744-6155.2006.00076.x.

Vinokur, A., Price, R., & Schul, Y. (1995). Impact of the JOBS intervention on unemployed workers varying in risk for depression. *American Journal of Community Psychology, 23*(1), 39–74. doi:10.1007/bf02506922.

Voogdt-Pruis, H. R., Beusmans, G. H., Gorgels, A. P., & van Ree, J. W. (2011). Experiences of doctors and nurses implementing nurse-delivered cardiovascular prevention in primary care: A qualitative study. *Journal of Advanced Nursing, 67*(8), 1758–1766. doi:10.1111/j.1365-2648.2011.05627.x.

Part I
Individual Physical, Cognitive, and Emotional Vulnerability Across the Life Course: Benchmarks and Developmental Challenges

Chapter 2
Benchmarks, Developmental Challenges, and Risks During the Prenatal and Infancy Period

Hellgard Rauh and Karl E. Bergmann

Introduction

The 1989 UN Convention on the Rights of the Child has not yet been signed and put into action by all nations. Nevertheless, the 2000 UN Millennium Declaration and the UN Millennium Goals for 2015 indicate that these initiatives appear to have stimulated a wealth of investigations and research, within nations and among nations, that aim at improving the lives of children, particularly those of very young children (Britto, Yoshikawa, & Boller, 2011; U.N. 1989, 2000, 2010). The foremost goals of the UN initiatives, with regard to young children, are reducing child mortality, improving maternal health, combating highly infectious diseases (such as HIV/AIDS and malaria), eradicating extreme poverty and hunger, achieving universal primary education, and promoting gender equality and empowerment of women particularly by addressing birth rate and children's development. The other major environmental and sociopolitical aims are ensuring environmental sustainability and a global partnership for development.

A Historical Sketch of Prevention in Early Childhood

Preventive public health care became linked to basic human rights already during the 1848 revolutions in Europe (Stöckel, 2007). These revolutions were stimulated by the democratic concepts of the French Revolution of 1789 and instigated by the consequences of rapid industrialization and urbanization with ensuing health problems in cities and disintegration of rural areas.

H. Rauh (✉)
Department of Psychologie, Universität Potsdam, Potsdam, Germany
e-mail: rauh@uni-potsdam.de; hellgard.rauh@gmx.de

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_2,

In 1870, one in four children in Germany and other industrializing countries did not survive their fifth birthday. Poverty, uncontrolled growth of compact urban areas, long workdays of both parents in factories and the ensuing neglect of the children, and unhygienic conditions contributed to this situation. The last decades of the nineteenth century saw the introduction of preventive measures, some of which continue today.

The first measures of prevention were taken at different system levels. At the local level, they ranged from systematic city planning with parks in residential areas, improved public transportation, municipal sewage systems, indoor plumbing, and community gardens for urban families. Socially committed factory owners improved workplaces and dwellings for their workers. Medical science flourished, and medical care improved, especially with the work of Nobel Prize winners such as Robert Koch and Ignaz Semmelweis. Bacteria of specific diseases were discovered, and vaccinations became possible in the following decades. Puerperal fever, which had led to the death of many women during childbirth especially in research-oriented hospitals, was drastically decreased by strict hygienic barriers between hospital care and pathology units. Prevention was also introduced at a behavioral or lifestyle level. The first books on infant care by Alfred Grotjahn appeared in 1912 (Stöckel, 2007). Auguste Victoria, the socially committed last German empress, founded the first prevention research clinic in Berlin in 1909 (Bergmann, Bergmann, Richter, Finke, & Dudenhausen, 2009) that was primarily devoted to improving and popularizing prenatal and infant care, especially breastfeeding and prevention of chronic illness. One of these preventive measures, originally introduced to combat tuberculosis and rickets, is still part of a typical German lifestyle: the family Sunday walk in the parks and forests.

These preventive measures at national and local levels were largely effective. The childhood mortality rate decreased to fewer than 100 per 1,000 children in 1930, in spite of war, hunger, and a disastrous economic recession; this rate is similar to many countries in Africa today. Currently, the infant mortality rate is below 0.6 % in the United States, below 0.45 % in the European Union, and below 0.25 % in Japan (CIA, 2012). Child and infant mortality rates continue to be major indicators for a country's health situation in underdeveloped and developing countries. Prevention measures in developing countries are generally similar to those in developed countries 100 years ago. UNICEF initiatives and support reduced the child and infant mortality rate worldwide from 1990 to 2006 by 24 % not only through improvements in hygiene (especially safe drinking water), vaccination against infectious illnesses, provision of additional vitamin A, and insecticide-treated mosquito nets but also by supporting breastfeeding efforts and providing community health services (UNICEF, 2012).

The prevention of childhood death continues to be an important goal at all levels: parents, communities, nations, and international organizations. The idea of prevention has, however, changed as the apparent needs and problems of children have changed given the greater diversity in family compositions, the improvements in health care and mental health care, and the advancement of child development research. Whereas prevention originally addressed protecting a young child from

physical dangers or from specific illnesses and diseases, prevention efforts in the 1960s and 1970s focused on preventing or compensating for social and psychological risks of the young child and recognizing early risk indicators of unfavorable outcomes in later childhood and adolescence (e.g., school dropout, juvenile delinquency and risky behavior, mental health problems). Preventive interventions combined physical/medical support with compensatory education, as in the Head Start programs in the United States, or focused on education and stimulation for all young children, as in the Sesame Street television series that was copied and adapted in many countries. While the TV series was available to anyone who was interested, education-oriented prevention was usually restricted to children with specific risks (e.g., children who were born prematurely, are disabled, or are from socially disadvantaged families). Numerous programs were developed and implemented. Well-controlled experimental studies as well as follow-up studies were conducted at research institutions, especially in the United States and Canada. Quasi-experimental longitudinal studies followed risk and comparison infants into later childhood and adolescence. Advancements in medical research and health provision, as well as the expansion of research and academic training in behavioral and social sciences and improvements in international scientific cooperation and exchange, stimulated large-scale studies also in other developed and industrialized countries (e.g., Europe, Australia, and Japan). In recent decades, globalization has accelerated in part due to rapidly changing communication technologies. Research cooperation with developing countries is growing. With support of the World Health Organization (WHO), international classification systems, such as the International Classification of Diseases, have been adopted worldwide. Supported by UNICEF, indices for a country's overall quality of life are being developed, such as the Human Development Index (which comprises gross domestic product, education, and life expectancy), to better compare countries, especially developing countries, and to measure improvements or declines (Bornstein et al., 2012). Improvements and declines may have many causes. Major changes in the living conditions of even the youngest children may offer opportunities but may also pose challenges and new risks, in developing as well as in developed countries. Research-based preventive interventions, even when they have proved to be effective, have to be adapted to these "historical" changes, and transfer of expertise to developing countries has to account for sociocultural and even physical differences.

Issues Related to Risks in Infancy and Early Childhood

After thousands of years of very similar upbringing and care for children in the infancy period, the infant ecology has changed dramatically in the past century and even more rapidly in the last few decades.

Intentional birth control (i.e., the prevention of an unwanted child) has eventually become pharmaceutically and medically accessible and legally accepted in most, except some of the poorest, countries since the 1960s. Birth control has led to

fewer children being born to increasingly older and better educated mothers. Unfertile couples and single women can now have their own biological children through in vitro fertilization, donor sperm, donor egg, and surrogate pregnancy. Although hotly discussed because of the ethical implications, methods of preimplantation genetic diagnostics, embryo selection, and genetic modification will be accessible in the near future.

Close surveillance of fetal gestation increases the number of pregnancies ending in live births, as well as the numbers of premature and multiple births. Adaptive changes in medical and psychological perinatal care have dramatically improved the survival chances of premature newborns and reduced major neurological, sensorimotor, and respiratory birth defects. Because of these advances, more *fragile neonates* and infants with overt or discreet health restrictions survive and have a good chance to live well into adulthood.

Family structures and family values, even within a culture or nation, have become extremely varied, and for many children they are unstable and unpredictable. Many children grow up in single-parent households or in nontraditional patchwork families (McCall, van IJzendoorn, Juffer, Groark, & Groza, 2011). The responsibility for a child can be split between the parents resulting in total setting changes for the infants on a daily or weekly basis, as sometimes in the case of separated parents. In modern urban settings, social and psychological care for the infant has become a complex and demanding task even in traditional families. Many infants may spend hours outside the home and with different caregivers several times per week or even per day. The number of directly accessible people for a young child is often highly restricted, and the contact periods are highly structured and only minimally controllable by the infant. New parents often have limited experience with young children and/or are skeptical about parenting styles of previous generations. They are left with being uncertain about parenting and are coping with parenting helplessness.

Child neglect, *child maltreatment*, *and even child abuse* have become public concerns involving children even at very young ages. Primary prevention focuses on introducing new parents to sensitive and secure parenting or on addressing factors that prevent parents from realizing the needs and reactions of their infant (e.g., stress, maternal depression).

Fewer and intentionally conceived children are more precious for their parents and receive more parental investment (Belsky, Steinberg, & Draper, 1991). *Overprotection and overstimulation* can become harmful for the developing child. Some parents feel pressure to start education as early as possible to prepare their children for the demands of a future world. Some children may even suffer under the load of parental expectations. Parents may cognitively overstimulate their babies with an overabundance of toys and introduce even toddlers to using media and electronic communication (e.g., Skype). It is still unknown how new electronic possibilities will influence or change the mental, communicative, and emotional development of preschool children.

Early institutional care is apparently inevitable not only for orphans but also for children of physically, mentally, socially, or emotionally dysfunctional parents

when there is no supportive network of relatives. In most developed countries, childcare institutions have become safe and child-oriented places, but they cannot replace the atmosphere of family upbringing. Therefore, foster care and adoption are generally preferred. In the 1970s, *adoption and foster parenthood* was advocated as being nearly equivalent to biological parenthood. Recent research is less optimistic, especially in the case of international adoption (Grotevant, 2011). Timing of adoption is important. Attachment research has demonstrated the sustaining importance of a first intensive socio-emotional bond between an infant and a primary caregiver. Even in the case of adoption or foster care shortly after birth, the infant's psychobiological system has already been shaped during pregnancy, as recent epigenetic research has shown (Meaney, 2010), for example, with regard to stress reactivity. Epigenetic research findings should be considered in surrogate situations, especially when the surrogate mother lives in poverty.

The following sections present overviews on recent scientific insights into typical development in the prenatal and perinatal periods, infancy, and toddlerhood, along with period-specific risks und vulnerabilities and prevention issues. For the postneonatal periods, and with reference to Bronfenbrenner's ecological model (Bronfenbrenner & Crouter, 1983), vulnerabilities are discussed relative to proximal and distal risks in age-specific environments and settings (i.e., for quality of parenting, quantity and quality of extra-parental care and education, institutional care, and poverty as an overarching impact).

Development, Risks, and Prevention in the Prenatal Period

Advanced methodology in biological, medical, and psychological research has contributed to increased knowledge of regular and deviant prenatal development and positive and detrimental influences during this period.

Throughout the 40 weeks of gestation, from fertilization until birth, the fetal organism develops not only by unfolding a genetic program. Fetal development is constantly stimulated, channeled, and controlled by processes from within the fetal organism, by the biological environment provided by the womb, by maternal metabolism and activity, and even by external influences. These influences interact and change with fetal organic and functional development. At birth, the child is just one of the possible developmental variants given the initial genetic endowment.

Most influences of the interacting prenatal milieu systems have a stimulating, channeling, and shaping effect and can be considered continuous adaptations of the fetal organism to its present situation. From an evolutionary perspective, some of the epigenetic effects could be interpreted as preadaptation to an external world as translated to the fetus by the maternal organism (Bornstein, 1989a; Fox, Levitt, & Nelson, 2010). Recent epigenetic research (Azar, 2011; Meaney, 2010; Zhang & Meaney, 2010) has revealed some of the microbiological mechanisms that produce and change the specifically adapted gene expression and gene transcription, especially during cell specialization. These processes involve triggering or silencing

certain regions of the genome (by methylation) in a manner that is specific for each cell type. Some cell structures are more prone to epigenetic alterations than others (Zhang & Meaney, 2010, p. 447). Cells involved in metabolic functioning and nerve cells appear to be among those that are sensitive to epigenetic reprogramming. These biological microprocesses explain the rather enduring effects of some early influences, although epigenetic reprogramming does not seem to be totally limited to the prenatal period.

For a long period, the embryo and fetus were considered largely protected from external influences except for exposure to alcohol, radiation, and specific maternal infections that had been proved to seriously disturb prenatal organic development especially in the first trimester. Epigenetic understanding will differentiate scientific knowledge of period-specific effects of non-noxious as well as noxious, or damaging, influences (teratogens).

In the *first 3 months* (first trimester) when most of the organic structures are being built, teratogens may cause a severe developmental deviation or even growth cessation of a specific inner organ or limb, as was sadly demonstrated in people prenatally exposed to thalidomide. In the *second trimester*, most of the brain architecture is being differentiated and functionally specialized (Fox et al., 2010). The hormonal milieu, for instance, as produced by the male infant and the mother, stimulates sexual differentiation of the organs but also of brain regions (Berenbaum & Snyder, 1995; Dörner 1989). Additional hormones introduced into the fetal–maternal system externally (e.g., by medication, polluted water, pesticides) can alter the developing neural sexual centers (Dörner et al., 2001). In the *last trimester*, the fetal system becomes more interconnected (e.g., between the sense organs and the brain), more differentiated, and partly adapted to smooth functioning by reduction of superfluous nerve cells and synapses. Infant gross and fine motor activity contributes to this process, and most senses are being tuned to the expected external world. The fetus becomes accustomed to basic melodic and rhythmic features of the mother, to acoustic-endocrine associations of acoustic emotional information, to different kinds of touch, and even to culture-specific tastes (Fox et al., 2010). Major detrimental conditions in this period can trigger premature birth. If born premature, the infant is now usually viable though still immature and highly vulnerable. More subtle stresses for the fetus in this period may affect the neuro-psycho-physiological functioning of the infant and may become apparent only as age-inappropriate imbalance of the physiological and/or the attentional systems of the newborn and the young infant.

The concepts of *sensitive and critical periods* have been frequently applied to obviously time-bound vulnerabilities or susceptibilities in prenatal, perinatal, and postnatal development (Bornstein, 1989a; Fox et al., 2010). These concepts could give the illusion that there is a special time of vulnerability or heightened susceptibility either for a range of functions or for specific types of outcomes. Research has, however, shown that the processes are much more varied, complex, and specific and that these concepts cannot, at present, serve for more than a rough and superficial orientation.

Several research activities have studied specific influences on fetal development that are potentially measureable and amenable to preventive actions. Some prominent examples are prenatal *exposure to external toxicants*, *maternal drug use*, and *maternal stress*.

Air, water, and soil, as well as commercial products, are replete with chemicals from factories, traffic, fertilizers, pesticides, sewage, and food additives not only in industrialized countries but also, and sometimes even more severely, in the developing countries. Free market practices usually allow the introduction of new chemicals without proof of their nontoxic effects on young organisms. Because of the unavoidable mixture of potential *toxicants*, it is difficult to prove the impact of a specific toxicant on the fetus and on the growing child. Koger, Schettler, and Weiss (2005) summarized and interpreted the existing empirical evidence. According to Koger and colleagues, toxicants interfere with brain development in critical periods, thereby affecting sensory, motor, and cognitive development. But they can also increase the propensity for aggression and violence in adolescence and adulthood. The former East Germany was late in prohibiting the use of DDT, giving room for a quasi-experimental design between cohorts from the former East Germany and West Germany. The study found that there was a significant difference in estrogen-sensitive female cancers in East Germany (Dörner, 1989).

In many countries, individuals are considered responsible for their own *drug use*. A partial exception is tobacco exposure that cannot be avoided even by strict nonsmokers. This might have been the reason why nations have been effective in prohibiting smoking in public places. Pregnant women are strongly advised to avoid smoking.

Deciding on the use or restriction of assumed *necessary medication* during pregnancy is much more difficult (e.g., in the case of maternal epilepsy). Even when controlling for maternal education and age, epileptic attacks during pregnancy, maternal postnatal infirmity, prenatal exposure to specific antiepileptic drugs, and combinations of drugs appeared to have a lasting detrimental effect on intelligence and information processing in offspring (Koch et al., 1999; Titze et al., 2008). This effect is acerbated when postnatal drug exposure is followed by, and combined with, social risks such as low-quality parental care and stimulation. A major challenge to the pharmaceutical industry is, therefore, to test for and acknowledge long-term behavioral consequences of prescribed drugs.

Alcohol is the most common substance used. Its teratogenic effect in the embryonic stage of prenatal development is well documented. The effects range from physical/facial signs and reduced brain volume and brain differentiation with ensuing severe mental retardation to impaired information processing and imbalanced emotional regulation. Elliot and colleagues (2008) reviewed epidemiological studies of prenatal alcohol exposure and fetal alcohol spectrum disorders (FASD) and experimental studies of primary and secondary prevention. They found that the most effective primary prevention of FASD was alcohol prohibition. There was no evidence that warning labels on alcohol containers or mass education campaigns reduce alcohol consumption in pregnant women. Individualized interventions were generally more effective in women who consumed low levels of alcohol at study

entry. High-risk women were less able to change their drinking behavior. The authors complained about the sometimes infrequent, and usually not coordinated, use of assessment instruments, the evaluation criteria or measures, and the study designs. They regretted the lack of evidence-based recommendations for effective prevention.

While toxicants, drugs, and alcohol can be considered objectively definable external influences, *maternal stress* is a subjective reaction to previous and current experiences. When the first studies of stress effects during pregnancy appeared, some women became even more anxious about the possible impact on their child, and others used potential stress effects as an argument for abortion. The effects of different kinds of stress have since been studied more carefully. These effects are generally not comparable with those of external influences, but they can add to the effects of external factors. Schetter (2011) reviewed the available studies. In a population with low socioeconomic status, maternal depressive symptoms and general distress predicted, with other potential influences controlled, growth retardation and low birth weight of the infant.

More *acute forms of stress* (e.g., death of a family member, earthquake, terrorist attack), chronic stress (e.g., general strain, household strain, homelessness), poverty, and neighborhood violence predicted more frequent premature birth. Good and effective partner relationship were positive, protective influences, but social support had varying effects depending on ethnicity, race, cultural values, and type of support. Severe chronic stress in late pregnancy was often associated with heightened infant irritability and emotional and attention-regulation problems in childhood (Schetter, 2011).

Population studies of those in the higher socioeconomic strata differentiate degrees of tolerable, or even *mild, stress* experiences. In a study by DiPietro and colleagues (2010), pregnant women were followed prospectively from 24 weeks gestation until the first weeks after birth. In this population, tolerable to mild stress appeared to have a slight accelerating effect on some infants, especially boys, with regard to motor maturity and neural conduction as a measure of information processing. They were also less irritable as neonates. In a previous study, infants born under similar conditions even had higher developmental test scores at age 2. Although severe and multiple stress can retard development or even terminate the gestational period by premature birth, some light or positive stress may benefit the infant's neural development in the perinatal period.

In the prenatal period, the fetus is not isolated from influences on its organism. These influences normally stimulate and channel the next developmental steps. They interact in a probabilistic way (Meaney, 2010; Rutter, 2000) and in a systemic manner (Sameroff, 2010) to build and form a unique individual, partially pre-adapted via the effects of maternal lifestyle and health status. Single environmental toxicants or noxious influences can usually be tolerated or compensated for by the developing fetus when they impinge in low dose and when there are no additional biological or social adversities, such as maternal drug use or abuse, poor nutrition, poor prenatal care, low maternal education, and poverty (Koger et al., 2005; Rutter, 2000).

When, however, several risks cumulate during gestation over longer periods, an extreme effect on the fetus could be the termination of growth and stillbirth; in less severe circumstances, the effect could be premature birth and low birth weight after intrauterine growth retardation. It is usually the fetus, not the maternal organism, that induces premature birth. The kind and intensity of external influence, the number and variety of additional previous and contemporaneous stressors, the actual developmental status of the fetus, and the specific individual vulnerabilities combine to determine the kind and extent of the effect at the microlevel. The effect will usually be viable, alternative developmental paths with either more or fewer degrees of freedom. The outcomes will usually not be defects but rather variants. In some extreme cases, these variants are not viable and lead to spontaneous abortion. In other cases, as with trisomy 21, the superfluous genetic information of the extra chromosome disturbs the prenatal fine-tuning of the developmental course, resulting in usually viable neonates but with individually very diverse combinations of extra problems (i.e., prenatal organic deviations). It seems that the extra load of genetic information leads to reduced epigenetic resilience (Rauh, 2006).

Prevention during pregnancy usually means primary prevention from the perspective of the child, primary prevention for some women, and secondary prevention for high-risk women. Positive social awareness of pregnancy may be an important factor in introducing preventive interventions when necessary. Pregnancy has become "chic" in Western countries, and young women overtly show their status instead of hiding it in large clothing as was the style in previous times. In many countries, women have regular, free gynecological checkups, take preparatory classes, and/or visit a midwife long before the birth. The expectant father is involved, and pregnancy-specific partner problems are discussed. Special help is offered, or even mandatory, if abortion is being considered, in situations of high stress, or in early diagnosis of a severely handicapped fetus. Some countries offer professional advice and support. All this has contributed to a low incidence of neonatal death rate in these countries. Abortion rate, especially for psychosocial reasons, should be considered a benchmark for a society's responsibility for its children. However, even in countries with a well-developed health and social system, the most-needy people fall through the net because they seem to be unable or unwilling to use the options available.

Risks and Prevention in the Perinatal Period

Maternal deaths during or shortly after delivery have become extremely rare in countries with high standards of prenatal, medical, and social care, even with increased maternal age at delivery. They are still an important issue in some developing countries. In more affluent countries, the mother–newborn dyad and optimizing early physical and socio-emotional care have come into focus.

The human infant is born immature neurologically and physically, even compared with other primates. Brain development extends well into the second year of

life and is not completed until adolescence. The human newborn's poor motor abilities have long camouflaged its differentiated competencies. Therefore, until recently, the neonate had been considered a brainstem-controlled, immature biological system. Psychological and biological research has since completely altered this image, and accordingly neonatal care.

The *newborn infant* is now considered a human being with typically human, albeit still immature, attributes such as perceptions, feelings, psychological needs, interests, basic knowledge, and social orientation. His or her biopsychosocial behavior is organized in labile systems (physiological, motor, tonus, arousal states, attention) (Als, 1986; Prechtl & Beintema, 1968); he or she shows active efforts of self-regulation, orientation, and interaction (Brazelton, 1973) and has astonishing acoustical and visual perceptual competencies, especially for social stimuli (Bremner & Fogel, 2001; Bremner & Slater, 2004; Fantz, 1961, Schaffer, 1989). *Competent Infant* (Stone, Smith, & Murphy, 1973) has become a catchword for infancy researchers and practitioners. Although being open to, and dependent on, socio-emotional interactions early on (Bowlby, 1958), the infant forms a strong emotional and person-specific tie to the primary caregiver not before the second half of the first year.

In most industrialized countries, it has become standard that the laboring woman can choose among several *delivery modes*, the child's father can participate in birth, the healthy newborn stays with the mother or parents immediately after birth before the general diagnostic procedures are applied, the infant's crib is right next to the mother during most of the day, and lactating and breastfeeding are strongly encouraged. Psychological needs of the baby and parents are respected, or even celebrated. Mutual familiarization of mother and father in the first hour after birth (Klaus & Kennell, 1976/1982) and extensively in the neonatal period appears to facilitate a strong emotional tie, or bond, of the parents to their child and promote sensitive parenting.

Newborn Infants at Risk

As one outcome of improved medical prenatal care, more newborns are considered at risk, or even at high risk, today than in previous decades.

The number of children who are born prematurely and/or with very low birth weight is increasing. This increase is due to older maternal ages at birth, close monitoring of potentially risky pregnancies, and an increasing number of multiple births following artificial facilitation of pregnancies. Maternal chronic diseases (e.g., epilepsy, AIDS, diabetes, kidney problems) or detrimental lifestyles (e.g., drug or alcohol use, poor nutrition) may also result in newborns at risk being underweight or overweight, sick, or addicted.

Neonatal care of at-risk infants has been successful over the past 50 years and a model for early intervention and prevention. Even extremely premature babies with little more than half the full-term gestational age of 40 weeks now have a fair

chance not only to survive but also to have fewer and less severe physical or neurological sequelae than ever before. Technological advances (e.g., incubators, mechanical respiratory support, refined intubations), pharmaceutical innovations (e.g., surfactants for accelerating lung maturation), and surgical refinements contributed to increased premature and neonatal survival, albeit at sometimes high costs of lasting physical and/or neurological functional damages (e.g., lungs, brains, or eyesight). Psychologically based sensitive observations led to suggesting psycho-physiologically better adapted neonatal environments and care (Als, Duffy, McAnulfy, & Badian, 1989; Hess, 2005; Schott, Broghammer, & Poets, 2011). Incubator noise was reduced, alarm beeps were replaced by other means to inform caregivers, a night–day rhythm was introduced, and painful procedures were replaced by less invasive methods; infants are being dressed and placed into nest-like cushions. With very fragile infants, stimulation has been reduced, including stroking the infant in situations of high arousal. Soft stimulation (auditory, visual, tactile), at least partly controllable by the infant (e.g., water beds), is being used with more stable infants. Mothers and fathers are not only allowed to see their child regularly but often included in their infant's care. "Kangarooing" (i.e., skin-to-skin relaxed contact of mother or father, a method first developed in poverty areas in Colombia) of very premature infants was first ridiculed in the high-tech neonatal intensive care units but has been accepted even in university intensive care wards as an effective way to improve the infant's breathing and temperature stability (Cattaneo et al., 1998). If positively accepted by parents, kangarooing can help them learn the subtle behavioral signals of their infant and fine-tune their psycho-physical communication with the fragile baby. Parents report higher confidence in infant care and appear to develop more sensitivity in the interaction with their baby (Chwo et al., 2002; WHO, 2003).

The *attitude change of parents and professionals* toward the newborn and premature baby (i.e., realizing the individual even in the neonate) has not only improved and refined neonatal and intensive care but also reduced some of the highly intrusive, and often painful, mechanical supports and monitoring, alleviated the initially strict regulations of care and visiting, and emphasized the parent–infant interaction as important for any newborn, but especially for infants at risk. Pediatric practice and developmental psychological research had participated in these changes.

Perinatal Risks and Long-Term Outcomes

With improved neonatal care, even very premature and sick newborns can survive. As a consequence, the number of people with physical handicaps, chronic illnesses, and physical vulnerability who will reach adulthood has risen significantly (Bergmann et al., 2009). Biological prenatal and perinatal risks continue to be major threats to an infant's physical, emotional, and social development. Premature birth, low birth weight, and early central nervous system injuries are still the main

risk factors for cognitive, motor, and language development; functional deficiencies; and compromised school achievement. Even when these children are functioning well in familiar and less complex situations, they can activate only reduced resources for complex problems (Evrard et al., 2011) and are more prone to psychosocial problems, psychopathology, and maladaptation.

An overview of longitudinal data confirmed several assumptions. (1) An increased incidence of *functional problems in later life* for children with prenatal risks can only be predicted at the statistical group level, not at the individual child level. Risk is a probabilistic concept based on group statistics (Farran & McKinney, 1986; Rutter, 2000). This is bad news if limited prevention resources are to be allocated to the neediest individuals. It is, however, good news against any early "negative selection," and it is also good news for parents hoping for good outcomes. (2) The same risk condition can have quite diverse outcomes emerging at different times in the life course. (3) A single risk factor—even a severe one—can often be compensated for or overcome. For all risk groups, this is most probable for children who experience a positive socio-emotional climate and interaction in their families in the early years. Low-quality parenting can aggravate the consequences of even minor risks. (4) Biological risks, such as premature birth and perinatal and postnatal illnesses, appear to affect motor, cognitive, and language development primarily in the first few years of life, whereas social risks (psychopathology of parent, inadequate parenting, inappropriate or lack of stimulation) become increasingly effective in the later years of development (Bornstein, 1989b; Carta et al., 2001; Laucht, Esser, & Schmidt, 2002; O'Connor & Rutter, 1996; Rauh, 1989, 2005). Exposure to chronic risk situations increases the probability of adverse outcome, often even in a curvilinear fashion. A combination of both biological and social risks tends to have an overadditive effect—that is, infants born at biological risks are even more susceptible to psychosocial and social risks and children in highly stressful environments are more prone to encountering biological and social risks (cascading effect; Koger et al., 2005).

In research and in practice, *risk indices* combining biological and social risks focus on perinatal risks. Prechtl (1980) suggested instead an "optimality index" assessing how close to completely healthy a newborn is. This concept and a mixture of risks and optimality seem to have entered indices for older children, as in the Child and Youth Well-Being Index (Foundation of Child Development, 2010).

Risks, Vulnerabilities, and Prevention in Infancy and Toddlerhood

Typical Development Benchmarks

In the *first 2 months*, the human infant is considered by many researchers to be primarily an extrauterine fetus adapting to the new physical and physiological

conditions, stabilizing the basic behavioral systems, practicing the senses, and differentiating the brain structures. A major "milestone" with great impact on the parents is the social smiling at about 6 weeks after expected birth date. The infant becomes increasingly active in social and emotional interaction, intensively studying familiar human faces. Rochat and Striano (1999) characterize the changes between 2 and 3 months as "revolution" or "social birth." Already within the first few months, infants differentiate between physical and social stimuli and soon tend to react differentially: they smile and babble to even a "still" face of a living human to reactivate it but rarely to a puppet (still-face paradigm), even one that had been moving previously. They habituate to repeated stimuli and prefer a familiar person to a strange person and familiar language, ethnic features, smells, and tastes.

In the *second half of the first year*, infants not only become motorically more independent and explorative (sitting, locomotor activity, grasping) but seem to develop an intuitive understanding of major physical principles (behavior of objects of different sizes and masses, weight, gravity, physical causality, object permanence, local relations) as well as of human intentional, or unintentional, actions. At around 8 months ("second social revolution" according to Tomasello, 1999), they start to enjoy games that involve taking turns, and signal by social referencing to their partner that they associate actions and action results with the actor. They can now not only differentiate between emotional expressions of their adult caregiver but also use his/her encouraging or warning facial expressions as information for their own actions. At approximately their first birthday, they start to actively orient the caregiver to objects or events of their interest, share their own emotional experience with their caregiver, request information or evaluation, or even playfully tease their caregiver. During the second half of the first year, the infant gets tuned to those patterns of stimuli and events that are meaningful and characteristic for his/her culture, and he/she sometimes becomes insensitive to those physical differences that do not carry cultural meaning (e.g., specific language sounds, forms, color shades).

At *18 months*, the infant has grown into a competent toddler, coping well in a here-and-now world at a sensorimotor level. He/she has developed a basic sense of psychophysical self as differentiated from others; can move about and act in a still limited physical world of objects, plants, and animals; can refer his/her actions and success or failure to the self as actor; and can actively participate, interact, and communicate (nonverbally) in a concrete social world of real and familiar people. He/she can infer at least some psychical experiences of others and differentiate them from his/her own experiences (distinguish between own feelings and wishes and those of others, between what the other person can see or has just experienced and what he/she sees and knows).

During the *second year of life*, new cognitive competence emerges and revolutionizes the toddler's worldview: the ability to represent. The child eventually acquires the ability to represent the objective world in gestures, words, drawings, and somewhat abstract mental representations, as well as the social world in symbolic play, imitation, and role-play, initially using concrete actions then increasingly mental representations. By 18 months, the child discovers the self as

a counterpart, visibly reflected in a mirror and internally represented as a mental object. Via language, even subjective feelings can become reified, controllable, and amenable to cultural influences. These representational worlds seem to be unlimited and may sometimes overwhelm the young child. Logical structuring helps the child to eventually tame the representations of the physical world, starting with prelogical abilities of classification and seriation and turning to concrete and formal operations in school children and adolescents (Piaget, 1952). The represented social worlds will eventually be structured by social and moral rules and values. Toward the end of the second year, children are usually eager to participate in their social setting. They imitate adults and other children, learn social scripts and social rules, and communicate orally. They become susceptible to cultural socialization and even eager to become socialized and participate in enculturation (Pauen & Rauh, 2008).

Attachment, the Infant–Parent Relationship

Infant development is closely linked to parenting. All infants need an *attachment relationship* to grow. Attachment develops over the first year of life from a general openness to social stimuli and interactions to a clear distinction between familiar and unfamiliar people. With the infant's advances in locomotion and with cognitive structuring of the social and physical world, the attachment figure becomes the emotional center for the infant. The intensity of attachment peaks at the end of the first year and continues into the second and third years of life; it protects the child from getting lost and from threats and dangers. Attachment also serves as a secure base for exploration in the physical and social world. The child develops a system of strategies and emotional experiences to cope with novelties and threats. At the representational level, the system of attachment relationships and coping strategies eventually forms an internal working model for social relations in general and the self in these relations (Bowlby, 1969; see Stevenson-Hinde & Verschueren, 2002). Parent–child relationships continue throughout life, and individuals' socio-emotional relationships with friends, partners, and later their own children have general similarities and can even be traced back to those first relationships.

Three *major strategies* characterize infants' primary attachments (Ainsworth, Blehar, Waters, & Wall, 1978; Crittenden, 2008): A securely attached infant will explore his/her surroundings freely in the presence of an "attachment figure," will freely express positive and negative feelings, and will turn to and seek security, comfort, and support from the attachment partner when feeling frightened, unsafe, and endangered. Infants with an insecure–avoidant attachment relationship also have strong emotional attachments, but they balance, or even curb, their need for closeness and comfort on the basis of the caregiver's reaction to the infant's negative emotions. Infants with an insecure–ambivalent attachment relationship intensify their emotional expression of distress to compel their attachment partner's total attention.

All three strategies are basically effective to get, and keep, the caregiver close. A secure strategy is prevalent in most societies and is most typical in affluent and stable families with well-educated mothers. The frequency of insecure strategies varies with cultural history and values and with the infant's early parenting experiences (Crittenden & Claussen, 2000).

A key factor for individual differences in attachment strategies, or attachment quality, in the second year of life is *maternal sensitivity*, especially in the first months of a child's life. Sensitivity to a young infant means prompt and reliable reaction to the infant's distress for establishing emotional security and to the infant's expressions as a basis for communication and for learning regularities; empathic interpretation of, and reaction to, the infant's needs; and emotional warmth in distress as well as in joyful situations. The ingredients of sensitivity may change in later infancy and toddlerhood: promptness becomes less important and emotional tenor and teacher qualities become more important. The concept of sensitivity is closely related to the concept of intuitive parenting (Papousek & Papousek, 1987) and to the concept of scaffolding (Vygotsky, 1978). Maternal sensitivity has proved to be predictive not only for attachment quality. Sensitive parenting appears to be the major protective factor for infants born at early biological risks and for irritable babies. Maternal sensitivity and quality of first attachment were predictive of social competence in preschool and school. Social relations, friendships, and even later romantic relationships seem to build on those early experiences, although direct predictions over long periods of life are moderated by many additional factors (Englund, Kuo, Puig, & Collings, 2011; Grossmann, 1988). Some transgenerational studies (Hautamäki, Hautamäki, Neuvonen, & Maliniemi-Piispanen, 2010) suggest that the parent's own socio-emotional security status translates into parental sensitivity with his/her infant and into infant attachment quality.

Even in the first few years of life, young children present a *wide range of individual differences*, not only in attachment strategies but also in rate of physical, cognitive, language, and social development; gender; temperament; and stress reactivity. Within the normative ranges, rate of development is not yet stable in early childhood, and prediction into school years, adolescence, or even adulthood is limited (Brim & Kagan, 1980; Kagan, Kearsley, & Zelazo, 1980). Nevertheless, rate of cognitive development and level of achievement at a particular age are frequent measures of intervention efficiency. Gender and temperamental differences, including stress reactivity, are considered major mediators and moderators of stressful experiences and coping efforts.

The Risk of Dysfunctional Early Parenting

The concepts of attachment and sensitive care have entered the fields of infant developmental psychopathology and infant psychiatry. Extreme experiences of the

infant in his/her early social relationships (e.g., with an unresponsive caregiver, an emotionally unavailable mother, or an unpredictable caregiver) may induce the child to develop nonnormative strategies to address immediate paramount needs for self-protection and attachment, but these strategies may eventually result in pathological attachment strategies (Crittenden, 2008). The infant stress system appears to be extremely sensitive to *significant caregiving perturbations* early in life resulting in heightened anxious behaviors and in cognitive biases to reinforce anxiety. Pathological strategies usually originate as situationally functional in the early caregiver context, but they limit the child's potential for emotional, social, communicative, and often cognitive development (Goodman, Quas, & Ogle, 2010). Frightening parental behavior, high insecurity in the primary setting, and genetic disposition can be the causes of deviant attachment relationships that predict a host of behavioral problems in childhood and even adolescence (Van IJzendoorn, Schuengel, & Bakermans-Kranenburg, 1999).

Genetic disposition and temperament can act as moderators of differential susceptibility. Infants with a highly negative temperament, usually indicative of future behavior problems, were less self-regulated when they were in unresponsive relationships, but more self-regulated than control children when in responsive relationships (Kim & Kochanska, 2012). Recent epigenetic research with rat puppies (Meaney, 2010) suggests not only that prenatal maternal stress experiences epigenetically program stress vulnerability of the fetus but that sensitive early care (licking in the case of rat mothers) may have a preventive/protective effect on stress vulnerability in the pup.

In extreme cases of dysfunctional parenting, the child has to be secured by communal authorities. *Programs to promote effective parenting* aim at preventing such drastic measures (Sandler, Schoenfelder, Wolchik, & MacKinnon, 2011). In home visiting programs, specially trained nurses guide first-time mothers from low-income homes in how to care for the infant physically and psychosocially. These programs have been established in many countries and usually report positive effects on the infant (developmental indices) and the mother (self-reliance) (Astuto & LaRue, 2009). A well-controlled study of a weekly early intervention for infants at age 6 months and a later intervention at 24–28 months found significant effects on maternal warmth and nurturance after the early intervention and more contingent maternal cognitive responsiveness and language input with the second intervention (Landry, Smith, Swank, & Guttentag, 2008). The toddlers were more cooperative, communicative, and engaged in mother–child interaction and book-reading when their mother had received the intervention during toddlerhood. Response to the intervention of very low-birth-weight infants was similar to those of term infants. The study shows that relatively short, but targeted, interventions at focal developmental times can be quite effective. There are, however, only few methodically sound studies in this field, and the kinds of interventions vary greatly as do the targeted families. Little or no information exists on families that do not volunteer in programs or that resent any intrusion for personal or political reasons.

Maternal Employment and Non-parental Care

Maternal out-of-home employment during the child's infancy has been a concern in many countries, particularly during the child's first year. *In most European countries*, paid maternal leave is generously allotted with 14–16 weeks of full pay, shared by the employer and the state, and partial pay is provided for up to 12 or more months. Reemployment is guaranteed for up to 3 years. In Northern European countries, paternal leave for 2 months is encouraged. Parents tend to enjoy this time and the relative financial security. Accordingly, most day-care provisions are for children older than 12 months.

The situation is very different *in the United States*. Maternity leave of 12 weeks is generally unpaid. Parents are responsible for organizing and financing child care. The variety of solutions found by U.S. parents is, therefore, unparalleled in other countries. Brooks-Gunn, Han, and Waldfogel (2010) analyzed data of the nationwide National Institute of Child Health and Human Development (NICHD) study of early childcare of more than 1,000 children born in 1991, from infancy to school age. They did not find any general effect of first-year maternal employment on children's cognitive, language, social, and emotional development. The most positive child outcomes were found in the group with mothers who were employed part time compared with full-time or unemployed mothers. In the group of children whose mothers were employed full time, teachers observed more externalizing problems when the children were 4, 5, or 7 years old. However, both part- and full-time employed mothers chose day-care provisions for their children that were of higher quality than those selected by stay-at-home mothers. Part-time working mothers tended to be more sensitive with their children and provided a more stimulating home environment than mothers of the other two groups. Finding a socially approved way to cope with both challenges—securing the family income and job and providing secure and stimulating care for the young infant—was a characteristic of U.S. mothers who were employed part time.

Paid maternity leave extends only to 16 weeks *in Switzerland*. There is no tradition of group day care for children younger than 3. A partly retrospective study by Averdijk and colleagues (2011) found more teacher-reported problem behavior in 7-year-old children who had experienced extended group day care.

In the United Kingdom, families get paid maternity leave for the first 3 months (Sylva et al., 2011). A British study on families, children, and childcare analyzed data on more than 1,000 families. Positive effects of nursery care were found for cognition. High-quality care was predictive of more positive cognitive development and of emotional regulation at 18 months. Children from ethnic minorities scored lower on all measures. Positive maternal care was a major predictor of positive infant development and included sensitivity, non-harshness, and opportunities for stimulation. These findings are compatible with those of Landry and colleagues (2011).

A *Norwegian national cohort* study analyzed more than 28,000 families (Bekkhus, Rutter, Maughan, & Borge, 2011). At 9 months, practically all infants

were still being raised at home. However, at 18 months, two-thirds attended some type of group day care. Higher parental education and family income predicted a higher probability of the toddler attending day care. The strongest predictors of child behavior problems at 36 months (i.e., for distress or aggressiveness) were the child's temperament at 6 months and family risks, with hardly any contribution of group day care. Extended paid maternity leave and high-quality group day care in this Scandinavian country seem to have resulted in positive situations for families and infants relative to 10 years previously, when day-care conditions resembled those in the United States today.

Some children may react to *childcare as a stressful experience*. Children in childcare exhibit higher cortisol levels, as indicators of stress, than at home, especially toddlers attending full day care. Daily schedules with long periods of both provider-directed structured activities and frequent full-group transitions were associated with cortisol increases. Girls tended to react with an increase in anxious–vigilant behavior, and boys reacted with angry–aggressive behavior. Inhibited children exhibited more social fearfulness. Children at the extremes of the temperament spectrum were more sensitive to variations in environmental conditions (Phillips, Fox, & Gunnar, 2011).

Neither in the United States (NICHD, 2001) nor in other countries (Ahnert, 2010; Rauh, Ziegenhain, Müller, & Wijnroks, 2000) did early day care generally affect the quality of the mother–child attachment relationship. However, infants of low-sensitive and low-responsive mothers who spent more than 10 h per week in low-quality multiple childcare arrangements tended to exhibit more frequent insecurity or disorganized attachments to their mothers and were less positively engaged with their mothers over the next 3 years. The way the infant becomes familiarized to the new setting (abruptly or gradually) tended to affect the relationship, but only in infants older than 11 months at day-care entry (i.e., when they are cognitively mature enough to relate two alternating primary settings) (Rauh et al., 2000).

Educational and Compensatory Preschool Experience

The largest study on the pedagogical *effects of day-care and early preschool attendance* has been conducted by the NICHD research group on Early Child Care and Youth Development in the United States. The study assessed cognitive, social, and emotional aspects of child development and controlled for many potential influences. The researchers found, at 54 months, general effects of quality and quantity of day-care and preschool attendance in cognitive and language skills but not in the social or emotional domain. Only the actual quality of childcare was related to child social competence and social behavior (NICHD, 2003a). At 6 years, maternal sensitivity remained the main predictor and influence on children's social skills, even when controlling for child competences at age 54 months (NICHD, 2003b). Early child education experiences can, however, pose risks for some

children (increased social problem behavior). Children with early negative emotionality or difficult temperament appear to be more affected by the quality of day care and preschool than children who are emotionally stable. Children living in poverty also appear to benefit more than others from high-quality day care; low-quality care is more detrimental to them than it is for other children. High-quality preschool experience can ameliorate social, cognitive, and emotional deficiencies in family care and stimulation but has minimal measurable effects in children from homes with positive environments (Phillips & Lowenstein, 2011).

Poverty and sociocultural disadvantage in otherwise well-to-do societies are often combined with dysfunctional parenting. Compensatory educational programs aim at disrupting the circle of disadvantage. Center-based early intervention programs focus primarily on the child, whereas home-based interventions target mainly the parents (the mother) and the parent–child interaction. Blok and colleagues (2005) compared the effectiveness of 19 scientifically controlled compensatory programs with educational orientations that had varying delivery modes in the United States, Europe, Israel, and Turkey. Some had started at or even before birth, and others as late as age 5. Early intervention had a moderate effect on the cognitive domain but not on the socio-emotional domain, and effects more or less faded during the next 10 years. For the cognitive domain, center-based programs were more effective than home-based programs, irrespective of other characteristics, and the combination of both was most effective when they included parenting skill support. No effects were related to the socio-emotional domain. Other variables such as age admitted to the program, program duration and intensity, or forms of parental and family support had no statistically relevant effects.

Evidence-based *early educational interventions with disabled children* and their families have yielded mixed results. They have been less positive than had been hoped for, and their effects have not lasted long (Farran, 2000). They were most effective in less disabled children and in promoting language development in young children with speech problems. A program-specific effect of generalized stimulation of development beyond the effect of attending day care at all could not be verified. Farran complained that many programs had either been developed for other groups of children and were then transferred to disabled children or were solely repair oriented.

Institutional Care, Foster Care, and Adoption

Hardly anybody would have believed that the disastrous experiences with institutionalized infants and young children after the two World Wars could be repeated even 50 years later. The Romanian orphans, victims of an insensitive political "experiment," presented all symptoms of a major *hospitalization syndrome*: grossly retarded physical and psychological development and depression. The second step of this "experiment" was international adoption of many of these children into very positive home environments with dedicated foster parents. The results are broadly

published (Audet & LeMare, 2011; Beckett et al., 2006; Hawk & McCall, 2011; McCall et al., 2011; McCall, van IJzendoorn, Juffer, Groark, & Groza, 2012; Nelson, 2007; O'Connor, Ruer, Beckett, Kreppner, & English/Romanian Adoptees Study Team, 2000; Rutter et al., 2010; Tarullo, Garvin, & Gunnar, 2011).

The Romanian orphans were physically and mentally underdeveloped when they arrived in their foster or adoptive homes in England, Canada, the Netherlands, and the United States usually independent of the duration of their institutionalization. Many caught up significantly in the first 2 years after adoption. Infants who had been 6 months and younger at arrival in their new families thrived best. Children who were older than 6 months at adoption also thrived but much less than the younger adoptees. After a "spurt" in physical and mental growth until age 4, their developmental impulse weakened and some physical and mental deficits remained in the majority of them.

Four probably *deprivation-related behavioral patterns* were observed and discussed by Rutter and colleagues (2010). They occurred in increased frequency in children with institutional care beyond the first 6 months of life and were usually observed in combinations at age 6 and older: quasi-autistic behavior and inability to relate socially, disinhibited attachment or indiscriminate attachment, inattention/overactivity, and cognitive impairment. There was a vast range of individual differences. Most worrying, however, was an increase of conduct and emotional disorders in the early school years and especially in adolescence. These disorders were most strongly related to inattentiveness/hyperactivity and an increased incidence of peer relationship problems. The researchers assume that the second half of the first year of life and the second year are particularly vulnerable periods for psychic neglect. This is usually the time when socio-emotional attachment culminates, receptive and expressive language start, and the neurological basis for executive control functions differentiates. Van IJzendoorn and colleagues (2011) are reluctant to accept a general post-institutionalization syndrome. They nevertheless state clearly that for young children the institutional setting is pathogenic and should be classified as a type of child maltreatment, particularly in the form of structural neglect (p. 26).

These findings present extremes and cannot be generalized to institutional care or transferred into *foster and adoptive families*, generally. Still, institutional care remains detrimental for infants compared with family care. Audet and LeMare (2011) conclude from existing research that infants with prolonged institutional care and severe deprivation thrive better with highly structured, even authoritarian, parenting as signaled by less frequent occurrence of inattention/overactivity disorders in school age. In contrast, very early adopted children and children with minimal deprivation experience would be even more prone for this disorder under an authoritarian parenting style. Foster and adoptive parents should get support in their parenting efforts to adapt their parenting to the special needs of their adoptive or foster children.

Studies of adopted children originating from ethnic groups other than those in Europe (e.g., East Asia, Middle America) or with less severe institutional deprivation than children in Romania report alternative post-adoptive behavioral problems

(e.g., more internalizing problems in girls adopted from China) (Cohen & Farnia, 2011).

Under extreme conditions of institutionalization, some children are more negatively affected than others. Indiscriminate attachment, or attachment disorganization, in the relation to the "favorite caregiver" was found only in institutionalized children without a *specific genotype* (long variant of the serotonin transporter gene 5HTT). Those with this genotype appeared to be protected from these adverse effects (Bakermans-Kranenburg, Dobrova-Krol, & van IJzendoorn, 2011). Tarullo and colleagues (2011) assume, based on EEG studies, that indiscriminately friendly behavior in internationally post-institutionalized adopted children is the effect of early chronic neural hypoactivation. The lack of social interaction with a primary caregiver may interfere with neural development; increases in one-to-one social interaction may ameliorate the effects.

In many countries, institutional care seems to be unavoidable in some children's lives, especially in countries stricken by war and epidemics (e.g., HIV). Prevention measures should consider the cultural resources and needs of young children to avoid institutionalized early care. If this is impossible, institutionalized care should be better adapted to the needs of young children.

Poverty and Social–Ecological Risks for the Infant

Poverty is a strong, overarching factor *in the child-rearing context*. According to research reviews reported by Duncan (2012), poverty (defined as low purchasing power) in the United States is strongly related to family stress, compromised parental and child health status, poor nutrition, parental depression, inappropriate socio-emotional and cognitive development of children, poor choices in regard to care and educational settings, inappropriate parenting styles, and so forth. Poverty is usually associated with multiple risks, such as family and housing insecurity, overcrowding, and lack of amenities (Schoon, Cheng, Jones, & Maughan, 2011). Poverty appears to have stronger effects on cognitive development, and family disruption seems to be more salient for emotional/social adjustment. Schoon and colleagues (2011) argue that it is necessary to identify how risks co-occur, to examine the role of their timing and duration, and to determine which risk factors prevail. Evidence from UK longitudinal studies suggests that duration of risk exposure, early risk (indicating cumulative processes), and sensitive periods are all significant.

When comparing developmental periods, *poverty in the prenatal and the infancy periods*—when the brain, stress reactivity, and immune functions develop—appears to be most detrimental (Duncan, 2012). Prenatal *stress* is mediated by the mother's physiology; postnatal stress is mediated by parental care and cognitive stimulation. According to the biological sensitivity theory (Blair et al., 2011), early experience shapes the stress response system to meet expected environments, with consequences for regulating behavior in those environments. Highly supportive and

unsupportive environments lead to elevated stress physiology. In unsupportive environments, this increase would not be well regulated and stress hormones would remain elevated, facilitating reactive and inflexible forms of behavior and cognition (p. 1980). In supportive environments, regulation of stress hormones would occur and facilitate reflective and flexible forms of behavior and cognition.

According to Cabrera and colleagues (2011), the effects of socioeconomic and educational risk factors on cognitive outcomes in toddlers are mediated mainly through *maternal sensitivity*, and the effects of these risk factors on toddlers' social behaviors are mainly transmitted through maternal sensitivity and father engagement. According to Schoon and colleagues (2011), affective and close parent–child interactions appear to be important factors in reducing the negative effects of family poverty on children's cognitive development and, in particular, on their behavioral outcomes. Therefore, some children show positive adjustment despite exposure to family poverty and/or instability (Schoon et al., 2011).

Mothers' distress influences the level of cognitive stimulation offered, which in turn influences children's development. Maternal distress following persistent hardship with reduced access to economic resources is associated with *reduced investment* in children between ages 9 months and 3 years in terms of cognitive stimulation and parent–child interaction, with negative consequences for these children's further development (Schoon et al., 2011).

Cognitive stimulation may reduce, but will not eliminate, the negative effects of family poverty. Until 1986, *compensatory intervention programs* targeting families in poverty usually focused on the child's cognitive and language stimulation. Families and family interaction were considered part of the child's problem. According to Gershoff and colleagues (2005), even the best programs could not "inoculate" children from the debilitating costs of chronic exposure to poverty-related stressors during later childhood because they did not take a contextual focus (housing conditions, income of families, employment of parents). Therefore, even initial positive effects faded out. Families may have become dependent on the services, and at project termination they were not sufficiently knowledgeable about how to obtain those services to which they and their children were entitled (Farran, 2000). Lowell and colleagues (2011) describe a recent child- and context-oriented intervention with family support services that includes home visiting programs and family resource centers. One of the positive effects was that participant families accessed 91 % of wanted services compared with 33 % in other families.

Child well-being indicators are population-based statistics that provide a sense of whether a group of children can enjoy a good quality of life (O'Hare, 2012). They include key dimensions such as health, education, and economic and material well-being and are used to rank national efforts and to measure global interventions. Most indices were developed in North America and were adapted to other countries. The Child Development Index (Foundation for Child Development, 2010) was expanded to the Child Well-Being Index and is used by UNICEF. The Human Development Index (Health, Education, Economy) is used by WHO.

A group of developmental scholars (Bornstein et al., 2012) analyzed internationally collected data of developing countries and supplemented the information

with more detailed data in subsamples. The researchers studied children up to age 5 in 28 countries on such factors as nutrition, parenting, kinds of discipline, exposure to violence, and quality of home environment to determine each country's socioeconomic environment for its children. Called the *Human Development Index* (HDI), it comprised gross domestic product (GDP, purchasing power parity in US dollars), education factors (literacy rate and school attendance), and life expectancy. According to the researchers, malnutrition was, and still is, a major problem in most underdeveloped countries (Arabi, Frongillo, Avula, & Mangasaryan, 2012). Approximately 35 % of the children younger than 5 suffer from malnutrition, and 11 % of the global disease burden can be attributed to malnutrition. Growth is stunted in 195 million children in these countries. Tackling stunted growth and iron and iodine deficiencies had previously been among the most effective childhood interventions Therefore, a high-priority UNICEF goal for 2015 is the reduction of hunger and child mortality. However, living in countries with high HDI did not always translate into better nutrition for infants and young children.

GDP was a major predictor for differences in the quality of housing and material resources in homes. GDP and education significantly predicted variation in formal learning resources. There was, however, no correlation of any indices with informal learning resources (Bradley & Putnick, 2012).

In all countries, 29 % of caregivers believed that physical punishment is necessary to rear a child properly. Nonviolent discipline was more frequently found in high-HDI countries, and severe violence was more often evident in low-HDI countries. Level of education was the main predictor of the country differences (Lansford & Deater-Deckard, 2012).

Caregivers in all countries provide more socio-emotional stimulation than cognitive stimulation. The strongest correlation was between GDP and both kinds of stimulation. Cognitive caregiving and stimulation were practiced more by parents in high-HDI countries (Bornstein & Putnick, 2012). High and low socio-emotional caregiving, however, could be found at all levels of HDI. Playing with the child and acknowledging the importance of play were unrelated to any of the subindices and to HDI. No HDI was related to playing.

Living in poverty can have *different meanings*. Worldwide applicable economic poverty indices can place a whole country into the poverty group, with some in the country living well but with the majority living on few economic resources. Or a country can be placed among the wealthy countries but have subgroups that are relative poor—even if their economic situation is much better than that of the average person in a generally poor country. Except for extreme poverty, relative poverty appears to affect the feelings of self-worth or acceptance in the respected society (Hurrelmann, 2006). Psychological interpretations of one's socioeconomic status appear to influence the upbringing and development of children.

Furthermore, the potentially negative *side effects of an intervention* should be kept in mind (Masten & Narayan, 2012). Interventions may disrupt or undermine naturally occurring resilience and recovery processes. There is a risk of imposing culturally or developmentally inappropriate interventions; families and parents may

experience them as intrusive and as debilitating their parental rights. The problems are particularly relevant in international transference of intervention projects.

From an *evolutionary perspective* on human development, even major deviations from the normative Western pattern are not necessarily pathological. Ellis and Bjorklund (2012) found that the first 5 years of a child's life are a sensitive period for adjusting biobehaviorally to the specific ecological context and developing strategies to cope with expected problems and stresses. In contexts of low stress, mildly stressing experiences may sensitize, or even buffer, a child for future stressful experiences. In contexts of high stress, vigilant, or even unemotional, reactions may be an optimal survival strategy. This perspective does not deny dysfunctional and deviant early parenting but offers a fresh perspective that is amenable to prevention strategies.

Conclusions

The early period of life appears to be particularly vulnerable to, as well as protected from, external influences. The person that the infant will become, is also greatly influenced by the life course he/she takes, his/her social network and cultural frame, and historical events and changes.

There is growing, but not yet sufficient, knowledge on early developmental and risk mechanisms. How, when, and where does environmental input get translated into the individual organism, and what mediates long-term effects (Rutter, 2000)? Recent genome analyses and epigenetic research, hormonal–physiological research in on stress effects, research on brain functioning, and longitudinal developmental studies have opened new venues. Close interdisciplinary cooperation offers great chances to improve our understanding of developmental and risk mechanisms and consequently generate more effective and ethical ways of preventing adverse developmental outcomes.

Predictability of adult's personality of physical and psychological adaptability, of adult productivity, social participation, and happiness from infancy data is extremely low. Nevertheless, early influences can contribute significantly, even when their effects are not immediately evident. Effects of genetic differences, as well as early influences, may appear only in later periods of life as a result of a "sleeper effect" or mediating influences that occur in a "cascading" manner. Noxious influences in this early period may result in specific defects, but more often they lead to only slight alterations of the organism, instigating slightly deviating developmental pathways as a form of adaptation to an environment as signaled by the influence.

In the prenatal and early postnatal periods, the structural and epigenetic basis for the individual is formed. Because brain development in humans extends far into the postnatal period, social and emotional experiences contribute to this process, shaping brain structure and function as well as the reactivity to particular kinds of situations/stimuli and stresses. At least from the second half of the first year, the

child's activities, initiatives, and active learning processes contribute to forming the individual's personality. The developing child is preparing to cope with a wide variety of influences.

Infants depend on reliable personal social–emotional care and on personal social interaction and communication for not only their development but also their survival (Bergmann & Bergmann, 2003). Their enormous learning capacity develops in a setting of familiar caregivers and peers. Social deprivation, especially past the first 6 months of life, can lead to lasting deviations in socio-emotional and cognitive functioning, delimiting their future flexibility in adaptation. Inadequate parenting that could be characterized as neglect, maltreatment, or abuse can lay the groundwork for lifelong problems in psychological and social functioning. Insufficient, or inadequate, cognitive stimulation can cause insufficient neurological and cognitive functioning. Consequently, interventions that support young parents in positive interaction with their infant and supportive or compensatory education from parents and/or in childcare centers have proved to have a long-lasting payoff.

Infants and young children differ in their sensitivity to external positive or negative influences because of genetic differences, health problems, early prenatal and postnatal stressful influences, or early difficult socio-emotional experiences. Young girls seem to be generally less vulnerable than are boys. Infants and young children with a more "difficult" temperament and who are less balanced and less self-regulated in their emotional reactions, as well as children with less reliable life conditions, seem to be more negatively affected by inadequate conditions in their homes or day-care settings and other kinds of problematic life conditions. However, they are also more positively affected by positive conditions and interventions than are emotionally balanced children. Both kinds of children—those who are emotionally stable, well raised, and well educated and those who have had difficult experiences—can grow into healthy and interesting people who have a positive impact during their lifetimes, if they can transform their developmental experiences into positive coping.

The early period of life is an optimal time for effective and economical preventive measures. These measures, however, should be geared to the individual child and his/her conditions. They can improve or stabilize a child's health; they can improve a child's cognitive development for later academic and vocational demands; they can help the child become and remain adaptive for coping with life circumstances and unexpected difficulties. A major task is to globalize early prevention efforts and adapt them to the respective cultures and conditions in a way that parents can raise their children to live in a complex and multifaceted world. The concepts and results presented in studies on early development, on early influences, and on early prevention provide a note of caution that well-off countries should not blindly rely on the hope that they can easily replenish their decreasing populations with children born and raised in poor and underdeveloped countries. Investment in the countries' own children as well as improvements in the developmental conditions of children around the world should be major aims of early prevention.

References

Ahnert, L. (2010). *Wie viel Mutter braucht ein Kind? Bindung–Bildung–Betreuung: Offentlich und privat [How much mother does an infant need? Attachment–education–care]*. Heidelberg: Spektrum.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Erlbaum.

Als, H. (1986). A synactive model of neonatal behavioral organization: Framework for the assessment of neurobehavioral development in the premature infant and for support of infants and parents in the neonatal intensive care environment. *Physical and Occupational Therapy in Pediatrics, 6*, 3–55.

Als, H., Duffy, F. H., McAnulfy, G. B., & Badian, N. (1989). Continuity of neurobehavioral functioning in preterm and full-term newborns. In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development* (pp. 3–28). Hillsdale, NJ: Erlbaum.

Arabi, M., Frongillo, E. A., Avula, R., & Mangasaryan, N. (2012). Infant and young child feeding in developing countries. *Child Development, 83*, 32–45.

Astuto, J., & LaRue, A. (2009). Home visitation and young children: An approach worth investing in? *SRCD Social Policy Report, 13*, 4.

Audet, K., & LeMare, L. (2011). Mitigating effects of the adoptive caregiving environment on inattention/overactivity in children adopted from Romanian orphanages. *International Journal of Behavioral Development, 35*, 107–115.

Averdijk, M., Besemer, S., Eisner, M., Bijleveld, C., & Ribeaud, D. (2011). The relationship between quantity, type, timing of extended childcare and child problem behavior in Switzerland. *European Journal of Developmental Psychology, 8*, 637–660.

Azar, B. (2011). The psychology of cells. *Monitor on Psychology, 42*(5), 32–35.

Bakermans-Kranenburg, M. J., Dobrova-Krol, N., & van IJzendoorn, M. (2011). Impact of institutional care on attachment disorganization and insecurity of Ukrainian preschoolers: Protective effect of the long variant of the serotonin transporter gene (5HTT). *International Journal of Behavioral Development, 36*, 2–10.

Beckett, C., Maughan, B., Rutter, M., Castle, J., Colvert, E., Groothues, C., et al. (2006). Do the effects of early severe deprivation on cognition persist into early adolescence? Findings from the English and Romanian Adoptees Study. *Child Development, 77*, 696–711.

Bekkhus, M., Rutter, M., Maughan, B., & Borge, A. I. H. (2011). The effects of group daycare in the context of paid maternal leave and high-quality provision. *European Journal of Developmental Psychology, 8*, 681–696.

Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development, 62*, 647–670.

Berenbaum, S. A., & Snyder, E. (1995). Early hormone influences on childhood sex-typed activity and playmate preferences: Implications for the development of sexual orientation. *Developmental Psychology, 31*, 13–42.

Bergmann, K. E., & Bergmann, R. (Eds.). (2003). *Health promotion and disease prevention in the family*. Berlin: de Gruyter.

Bergmann, K. E., Bergmann, R. L., Richter, R., Finke, C., & Dudenhausen, J. W. (2009). Frühe gesundheitsförderung und Prävention am Beginn des 20. und 21. Jahrhunderts [Early health promotion and prevention at the beginning of the 20th and 21st centuries: Data and thoughts on the 100th anniversary of the opening of the Empress Auguste Victoria House in Berlin]. *Gesundheitswesen, 71*, 709–721.

Blair, C., Granger, D. A., Willoughby, M., Mills-Koonce, R., Cox, M., Greenberg, M. T., et al. (2011). Salivary cortisol mediates effects of poverty and parenting on executive functions in early childhood. *Child Development, 82*, 1970–1984.

Blok, H., Fukkink, R. G., Gebhardt, E., & Leseman, P. P. M. (2005). The relevance of delivery mode and other programme characteristics for the effectiveness of early childhood intervention. *International Journal of Behavioral Development, 29*, 35–47.

Bornstein, M. H. (1989a). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychological Bulletin, 105*, 179–197.

Bornstein, M. H. (1989b). Stability in early mental development: From attention and information processing in infancy to language and cognition in childhood. In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development* (pp. 147–170). Hillsdale, NJ: Erlbaum.

Bornstein, M. H., Britto, P. R., Nonoyama-Tarumi, Y., Ota, Y., Petrovic, O., & Putnick, D. L. (2012). Child development in developing countries: Introduction and methods. *Child Development, 83*, 16–31.

Bornstein, M. H., & Putnick, D. L. (2012). Cognitive and socioemotional caregiving in developing countries. *Child Development, 83*, 46–61.

Bowlby, J. (1958). The nature of the child's tie to his mother. *International Journal of Psycho-Analysis, 39*, 1–23.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York, NY: Basic.

Bradley, R. H., & Putnick, D. L. (2012). Housing quality and access to material and learning resources within the home environment in developing countries. *Child Development, 83*, 76–91.

Brazelton, T. B. (1973). *Neonatal behavioral assessment scale* (Clinics in Developmental Medicine, Vol. 50). London: W. Heinemann.

Bremner, G., & Fogel, A. (Eds.). (2001). *Blackwell handbook of infant development*. Oxford: Blackwell.

Bremner, G., & Slater, A. (Eds.). (2004). *Theories of infant development*. Oxford: Blackwell.

Brim, O. G., & Kagan, J. (1980). Constancy and change: A view of the issues. In O. G. Brim Jr. & J. Kagan (Eds.), *Constancy and change in human development* (pp. 1–25). Cambridge, MA: Harvard University Press.

Britto, P. R., Yoshikawa, H., & Boller, K. (2011). Quality of early childhood development programs in global contexts: Rationale for investment, conceptual framework and implications for equity. *SRCD Social Policy Report, 25*(2), 1–31.

Bronfenbrenner, U., & Crouter, A. C. (1983). The evolution of environmental models in developmental research. In P. H. Mussen (Series Ed.) & W. Kessen (Vol. Ed.) *Handbook of child psychology, Vol I: History, theory, and methods* (pp. 357–414). New York, NY: Wiley.

Brooks-Gunn, J., Han, W., & Waldfogel, J. (2010). First-year employment and child development in the first 7 years. *Monographs of the Society for Research in Child Development, 75*(2), 50–58. Serial No. 296.

Cabrera, N. J., Fagan, J., Wight, V., & Schadler, C. (2011). Influence of mother, father, and child risk on parenting and children's cognitive and social behaviors. *Child Development, 82*, 1985–2005.

Carta, J. J., Atwater, J., Greenwood, C. R., McConnell, S. R., McEvoy, M. A., & Williams, R. (2001). Effects of cumulative prenatal substance exposure and environmental risks on children's developmental trajectories. *Journal of Clinical Child Psychology, 30*, 327–337.

Cattaneo, A., Davanzo, R., Worku, B., Surjono, A., Echeverria, M., Bedri, A., et al. (1998). Kangaroo mother care for low birthweight infants: A randomized controlled trial in different settings. *Acta Paediatrica, 87*, 976–985.

Chwo, M. J., Anderson, G. C., Good, M., Dowling, D. A., Shiau, S. H., & Chu, D. M. (2002). A randomized controlled trial of early kangaroo care for preterm infants: Effects on temperature, weight, behaviour, and acuity. *Journal of Nursing Research, 10*, 129–142.

CIA. (2012). World factbook. Retrieved from http://cia.gov/library/publications/the-world-factbook

Cohen, N. J., & Farnia, F. (2011). Socio-emotional adjustment and attachment in children adopted from China: Processes and predictors of change. *International Journal of Behavioral Development, 35*, 67–77.

Crittenden, P. M., & Claussen, A. H. (Eds.). (2000). *The organization of attachment relationship. Maaturation, culture, and context*. Cambridge: Cambridge University Press.

Crittenden, P. M. (2008). *Raising parents: Attachment, parenting and child safety*. Portland, OR: Willan Publishing.

DiPietro, J. A., Kivlighan, K. T., Costigan, K. A., Rubin, S. E., Shiffler, D. E., Henderson, J. L., et al. (2010). Prenatal antecedents of newborn neurological maturation. *Child Development, 81*, 115–130.

Dörner, G. (1989). Significance of hormones and neurotransmitters in pre- and early postnatal life for human ontogenesis. *International Journal of Prenatal and Perinatal Psychology and Medicine, 1*, 145–150.

Dörner, G., Götz, F., Rohde, W., Plagemann, A., Lindner, R., Peters, H., et al. (2001). Genetic and epigenetic effects on sexual brain organization mediated by sex hormones. *Neuroendocrinology Letters, 22*, 403–409.

Duncan, G. J. (2012). Give us this day our daily bread. *Child Development, 83*, 6–15.

Elliot, L., Coleman, K., Suebwongpat, A., & Norris, S. (2008). Fetal alcohol spectrum disorders (FASD): Systematic reviews of prevention, diagnosis and management. *Health Services Assessment Collaboration* (HSAC) Report, 1(9)

Ellis, B. J., & Bjorklund, D. F. (2012). Beyond mental health: An evolutionary analysis of development under risky and supportive environmental conditions—An introduction to the Special Section. *Developmental Psychology, 48*, 591–597.

Englund, M. M., Kuo, S. I.-C., Puig, S., & Collings, W. A. (2011). Early roots of adult competence: The significance of close relationships from infancy to early adulthood. *International Journal of Behavioral Development, 35*, 490–496.

Evrard, D., Charollais, A., Marret, S., Radi, S., Rezrazi, A., & Mellier, D. (2011). Cognitive and emotional regulation developmental issues in preterm infants 12 and 24 months after birth. *European Journal of Developmental Psychology, 8*, 171–184.

Fantz, R. L. (1961). The origin of form perception. *Scientific American, 204*, 258–271.

Farran, D. C. (2000). Another decade of intervention for children who are low income or disabled: What do we know? In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (2nd ed., pp. 510–547). Cambridge: Cambridge University Press.

Farran, D. C., & McKinney, J. D. (Eds.). (1986). *Risk in intellectual and psychosocial development*. New York, NY: Academic Press.

Foundation for Child Development. (2010). *Child and youth well-being index*. Durham, NC: Duke University.

Fox, S. E., Levitt, P., & Nelson, C. A., III. (2010). How the timing and quality of early experience influence the development of brain architecture. *Child Development, 81*, 28–40.

Gershoff, E. T., Aber, J. L., & Raver, C. C. (2005). Child poverty in the United States: An evidence-based conceptual framework for programs and policies. In R. M. Lerner, F. Jacobs, & D. Wertlieb (Eds.), *Applied developmental science: An advanced textbook*. Thousand Oaks, CA: Sage Publications.

Goodman, G. S., Quas, J. A., & Ogle, C. M. (2010). Child maltreatment and memory. *Annual Review of Psychology, 61*, 325–351.

Grossmann, K. E. (1988). Longitudinal and systemic approaches in the study of biological high- and low-risk groups. In M. Rutter (Ed.), *Studies of psychosocial risk. The power of longitudinal data* (pp. 138–157). Cambridge: Cambridge University Press.

Grotevant, H. D. (2011). Emerging issues in bridging developmental science, practice, and policy in the best interest of children without permanent parents. In R. B. McCall, M. H. van IJzendoorn, F. Juffer, C. J. Groark, & V. K. Groza (Eds.), *Children without permanent parents: Research, practice, and policy* (pp. 274–281). *Society for Research in child Development Monographs*, *76*(4). Serial No. 301.

Hautamäki, A., Hautamäki, L., Neuvonen, L., & Maliniemi-Piispanen, S. (2010). Transmission of attachment across three generations. *European Journal of Developmental Psychology, 7*, 618–634.

Hawk, B. N., & McCall, R. B. (2011). Specific extreme behaviors of post-institutionalized Russian adoptees. *Developmental Psychology, 47*, 732–738.

Hess, C. R. (2005). NICU-based interventions for high-risk infants. In D. M. Teti (Ed.), *Handbook of research methods in developmental science* (pp. 295–316). Oxford: Blackwell.

Hurrelmann, K. (2006). *Gesundheitssoziologie [Sociology of health]*. Weinheim, Germany: Beltz.

Kagan, J., Kearsley, R. B., & Zelazo, P. R. (1980). *Infancy: Its place in human development*. Cambridge, MA: Harvard University Press.

Kim, S., & Kochanska, G. (2012). Child temperament moderates effects of parent-child mutuality on self-regulation: A relationship-based path for emotionally negative infants. *Child Development, 83*, 1275–1289.

Klaus, M. H., & Kennell, J. H. (1976/1982). *Parent-infant bonding*. St. Louis, MO: C. V. Mosby Company.

Koch, S., Titze, K., Zimmermann, R. B., Schröder, M., Lehmkuhl, U., & Rauh, H. (1999). Longterm neurological consequences of maternal epilepsy and anticonvulsant treatment during pregnancy for school-age children and adolescents. *Epilepsia, 40*, 1237–1243.

Koger, S. M., Schettler, R., & Weiss, B. (2005). Environmental toxicants and developmental disabilities. *American Psychologist, 60*, 243–255.

Landry, S. H., Smith, K. E., Swank, Z. T., Crawford, A. D., & Solari, E. F. (2011). The effects of responsive parenting intervention on parent-child interaction during shared book reading. *Developmental Psychology, 44*, 1335–1353.

Landry, S. H., Smith, K. E., Swank, P. R., & Guttentag, C. (2008). A responsive parenting intervention: The optimal timing across early childhood for impacting maternal behaviors and child outcomes. *Developmental Psychology, 44*, 1335–1353.

Lansford, J., & Deater-Deckard, K. (2012). Childrearing discipline and violence in developing countries. *Child Development, 83*, 62–75.

Laucht, M., Esser, G., & Schmidt, M. H. (2002). Vulnerability and resilience in the development of children at risk: The role of early mother-child interaction. *Revista de Psiquiatria Clinica, 29* (1), 20–27.

Lowell, D. I., Carter, A. S., Godoy, L., Paulicin, B., & Briggs-Gowan, M. J. (2011). A randomized controlled trial of Child FIRST: A comprehensive home-based intervention translating research into early childhood practice. *Child Development, 82*, 193–208.

Masten, A. S., & Narayan, A. J. (2012). Child development in the context of disaster, war, and terrorism: Pathways of risk and resilience. *Annual Review of Psychology, 63*, 227–257.

McCall, R., van IJzendoorn, M., Juffer, F., Groark, C. J., & Groza, V. K. (2011). Children without permanent parents: Research, practice, and policy. *Monographs of the Society for Research in Child Development, 76*(4), 223–272. Serial No. 301.

McCall, R., van IJzendoorn, M., Juffer, F., Groark, C. J., & Groza, V. K. (2012). The development and care of institutionally reared children. *Child Development Perspectives, 6*, 174–180.

Meaney, M. J. (2010). Epigenetics and the biological definition of gene x environment interactions. *Child Development, 81*, 41–79.

Nelson, C. A. (2007). A neurobiological perspective on early human deprivation. *Child Development Perspectives, 1*, 13–18.

NICHD Early Child Care Research Network. (2001). Child care and family predictors of preschool attachment and stability from infancy. *Developmental Psychology, 37*, 847–862.

NICHD Early Child Care Research Network. (2003a). Does quality of child care affect child outcomes at 4 1/2? *Developmental Psychology, 39*, 451–469.

NICHD Early Child Care Research Network. (2003b). Social functioning in first grade: Associations with earlier home and child care predictors and with current classroom experiences. *Child Development, 74*, 1639–1662.

O'Connor, T. G., Rutter, M., Beckett, C., Kreppner, L., & English and Romanian Adoptees Study Team. (2000). The effects of global severe privation and cognitive competence: Extension and longitudinal follow-up. *Child Development, 71*, 376–390.

O'Connor, T. G., & Rutter, M. (1996). Risk mechanisms in development: Some conceptual and methodological considerations. *Developmental Psychology, 32*, 787–795.

O'Hare, W. P. (2012). Development of the child indicator movement in the United States. *Child Development Perspectives, 6*, 79–84.

Papousek, H., & Papousek, M. (1987). Intuitive parenting: A dialectic counterpart to the infant's integrative competence. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 669–720). New York, NY: Wiley.

Pauen, S., & Rauh, H. (2008). Frühe Kindheit: Das Säuglingsalter [Early childhood: Infancy]. In M. Hasselhorn & R. K. Silbereisen (Eds.), *Enzyklopädie der Psychologie [Encyclopedia of psychology], Serie V Band 4: Entwicklungspsychologie des Säuglings- und Kindesalters [Developmental psychology of infancy and early childhood]* (pp. 67–126). Göttingen, Germany: Hogrefe.

Phillips, D. A., Fox, N. A., & Gunnar, M. R. (2011). Same place, different experiences: Bringing individual differences to research in child care. *Child Development Perspectives, 5*, 44–49.

Phillips, D. A., & Lowenstein, A. E. (2011). Early care, education, and child development. *Annual Review of Psychology, 62*, 483–500.

Piaget, J. (1952). *The origins of intelligence*. New York, NY: International Universities Press.

Prechtl, H. F. R. (1980). The optimality concept. *Early Human Development, 4*, 201–205.

Prechtl, H. F. R., & Beintema, J. (1968). *The neurological examination of the full-term newborn infant* (Clinics in developmental medicine, Vol. 28). London: S.I.M.P. with Heinemann Medical.

Rauh, H. (1989). The meaning of risk and protective factors in infancy. *European Journal of Psychology of Education, 4*, 161–173.

Rauh, H. (2005). The "at-risk" concept. In B. Hopkins (Ed.), *The Cambridge encyclopedia of child development* (pp. 393–397). Cambridge, UK: Cambridge University Press.

Rauh, H. (2006). Plastizität und Umweltbedingungen am Beispiel der Trisomie 21 [Plasticity and environmental influences in the case of trisomy 21]. In T. Klauss (Ed.), *Geistige Behinderung und Psychologische Perspektiven [Mental deficiency and psychological perspectives]* (pp. 45–72). Heidelberg, Germany: Universitätsverlag Winter.

Rauh, H., Ziegenhain, U., Müller, B., & Wijnroks, L. (2000). Stability and change in infant-mother attachment in the second year of life: Relations to parenting quality and varying degrees of daycare experience. In P. K. Crittenden & A. H. Claussen (Eds.), *The organisation of attachment relationships: Maturation, culture and context* (pp. 251–276). New York, NY: Cambridge University Press.

Rochat, P., & Striano, T. (1999). Social-cognitive development in the first year. In P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life* (pp. 3–34). Mahwah, NJ: Erlbaum.

Rutter, M. (2000). Resilience reconsidered: Conceptual considerations, empirical findings, and policy implications. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (pp. 651–682). Cambridge, UK: Cambridge University Press.

Rutter, M., Sonuga-Barke, E., Beckett, C., Castle, J., Kreppner, J., Kunsta, R., et al. (2010). Deprivation-specific psychological patterns: Effects of institutional deprivation. *Monographs of the Society for Research in Child Development, 75*(1), 232–247. Serial No. 295.

Sameroff, A. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Development, 81*, 6–22.

Sandler, I. N., Schoenfelder, E. N., Wolchik, S. A., & MacKinnon, D. P. (2011). Long-term impact of prevention programs to promote effective parenting: Lasting effects but uncertain processes. *Annual Review of Psychology, 62*, 299–329.

Schaffer, H. R. (1989). Early social development. In A. Slater & G. Bremner (Eds.), *Infant development* (pp. 189–210). London: Erlbaum.

Schetter, C. D. (2011). Psychological science on pregnancy: Stress processes, biosocial models, and emerging research issues. *Annual Review of Psychology, 62*, 531–558.

Schoon, I., Cheng, H., Jones, E., & Maughan, B. (2011). Resilience in children facing poverty and family disadvantage: Evidence from the UK millennium cohort. *International Society for the Study of Behavioural Development*. Bulletin 1, 5–8. Serial No. 59.

Schott, C., Broghammer, N., & Poets, C. (2011). Fortschritte in der Neonatologie—Vom Sicherstellen des Überlebens hin zu einer guten Lebensqualität [Advances in neonatology—From securing survival to securing life quality]. *Frühförderung Interdisziplinär, 30*, 128–136.

Stevenson-Hinde, J., & Verschueren, K. (2002). Attachment in childhood. In P. K. Smith & C. H. Craig (Eds.), *Blackwell handbook of childhood development* (pp. 182–204). Oxford: Blackwell.

Stöckel, S. (2007). Geschichte der Prävention und Gesundheitsförderung [History of prevention and health promotion]. In K. Hurrelmann, T. Klotz, & J. Haisch (Eds.), *Lehrbuch Prävention und Gesundheitsförderung [Textbook on prevention and health promotion]* (2nd ed., pp. 21–30). Bern: Hans Huber.

Stone, J. L., Smith, H. T., & Murphy, L. B. (Eds.). (1973). *The competent infant: Research and commentary*. New York, NY: Basic Books.

Sylva, K., Stein, A., Leach, P., Barnes, J., Malmberg, L.-E., & FCCC team. (2011). Effects of early child-care on cognition, language, and task-related behaviors at 18 months: An English study. *British Journal of Developmental Psychology, 29*, 18–45.

Tarullo, A. R., Garvin, M. C., & Gunnar, M. R. (2011). Atypical EEG power correlates with indiscriminately friendly behavior in internationally adopted children. *Developmental Psychology, 47*, 417–431.

Titze, K., Koch, S., Helge, H., Lehmkuhl, U., Rauh, H., & Steinhausen, H.-C. (2008). Prenatal and familial risks of children born by epileptic mothers: Long term effects on the cognitive development. *Developmental Medicine & Child Neurology, 50*, 117–122.

Tomasello, M. (1999). Social cognition before the revolution. In P. Rochat (Ed.), *Early social cognition* (pp. 301–314). Mahwah, NJ: Erlbaum.

U.N. (1989). Convention on the rights of the child. Retrieved from http://www2.ohchr.org/english/law/crc.htm

U.N. (2000). Millennium declaration. Retrieved from http://www.un.org/millennium/declaration/ares552e.htm

U.N. (2010) The millennium development goals. Retrieved from http://www.undp.org/content/undp/en/home/mdgoverview.html

UNICEF Press Information Service. (2012). Infant mortality rate.

Van IJzendoorn, M. H., Palacios, J., Sonuga-Barke, E. J. S., Gunnar, M. R., Vorria, P., McCall, R. B., et al. (2011). Children in institutional care: Delayed development and resilience. *Monographs of the Society for Research in Child Development, 76*(4), 8–30. Serial No. 301.

Van IJzendoorn, M. H., Schuengel, C., & Bakermans-Kranenburg, M. J. (1999). Disorganized attachment in early childhood: Meta-analysis of precursors, concomitants, and sequelae. *Development and Psychopathology, 11*, 225–249.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

WHO. (2003). *Kangaroo mother care. A practical guide: Department of reproductive health and research*. Geneva: WHO.

Zhang, T.-Y., & Meaney, M. J. (2010). Epigenetics and the environmental regulation of the genome and its function. *Annual Review of Psychology, 61*, 439–466.

Chapter 3
Childhood and the Entry into Adolescence: A Pivotal Period in Health-Related Behaviors and Prevention

Julia A. Graber, Julie C. Hill, and Mary E. Saczawa

Introduction

> Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity (World Health Organization, 1947, p. 100).

Across definitions, a critical component of health is that it is more than just the absence of disease. At the same time, it has been much easier to define the experience of disease than it is the experience of "physical, mental, and social well-being." Hence, most industrialized societies use systems for tracking morbidity and mortality as indicators of population health. When population data are parsed by demographic characteristics like age, gender, and ethnicity, childhood and the entry into adolescence are periods of relative good health with low mortality in comparison to younger and older ages. Despite the absence of serious disease and death, the entry into adolescence has been the target of numerous health-focused preventive interventions (Millstein, Petersen, & Nightingale, 1993; Ozer & Irwin, 2009). The rationale behind these initiatives is not based on current health per se as much as it is based on the critical aspects of development that occur as children become adolescents.

Understanding why health is an important concept to foster during the transition from childhood to adolescence requires understanding biological, psychological, and social aspects of development during this period. In particular, the transition into adolescence is defined by the biological and social–psychological experience of puberty; neuro-cognitive and emotional development; changes in social relationships with parents, family members, friends, and peers; and emergence of romantic interest. It is these developmental processes that make this transition pivotal for health.

J.A. Graber (✉)
Department of Psychology, University of Florida, PO Box 112250, Gainesville, FL 32611-2250, USA
e-mail: jagraber@ufl.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_3,

Whereas puberty is often identified as a central experience of adolescence, many individuals, especially girls, will begin puberty during the childhood years with pubertal development spanning much of the early part of adolescence (for a recent review of pubertal development, see DeRose, Graber, & Brooks-Gunn, 2010). In addition, at least in Western cultures that value independence, it is expected that during the course of adolescence, individuals will learn to become self-reliant and autonomous thinkers and decision makers. The comparison of autonomy and independent decision making between childhood and adolescence highlights the salience of these behaviors to health. During childhood, parents or caregivers are primarily responsible for children's health. It is expected that caregivers arrange for preventive health care and treatment as indicated. In addition, parents control a range of health-related behaviors such as selecting foods and activities as well as modeling healthy behaviors for children. Notably, surveillance and monitoring of population health to track mortality, morbidity, and conditions identified as precursors to disease, from prenatal development through childhood, are reliant on medical and government records and reports from parents and physicians about specific child behaviors.

In contrast, adolescents increasingly spend less time with parents and make more independent decisions about behaviors that in turn will influence their current mortality and morbidity. These decisions and behaviors (such as food choice or physical activity participation) are hypothesized to have lasting effects on health trajectories not only during adolescence but also into adulthood and across the lifespan. Because of this increased autonomy during the course of adolescence, national health surveys monitor the reports of adolescents themselves rather than parents. Furthermore, preventive interventions more often target adolescents directly and only in some cases include parents as part of health promotion for adolescents.

In this chapter, we discuss the emergence of health-related behaviors from a developmental perspective focusing on models of adolescent development and the contexts of development that influence health behaviors and choices. We begin with a brief review of mortality statistics and national health trends in the United States. We then review models of development that are particularly useful for understanding pathways to better versus poorer health outcomes. Finally, we discuss future directions in research and prevention.

Tracking Child and Adolescent Health

When looking at rates of mortality in the US population, as noted, childhood and adolescence are periods of low mortality. Age trends in rates of mortality are shown in Fig. 3.1; these data are for estimated deaths per 100,000 individuals in 2007 (Centers for Disease Control and Prevention [CDC], 2010a). Mortality rates are the highest among adults 65 and older with adults aged 45–64 having the next highest mortality rates. Apart from infancy, childhood (aged 5–9) and early adolescence

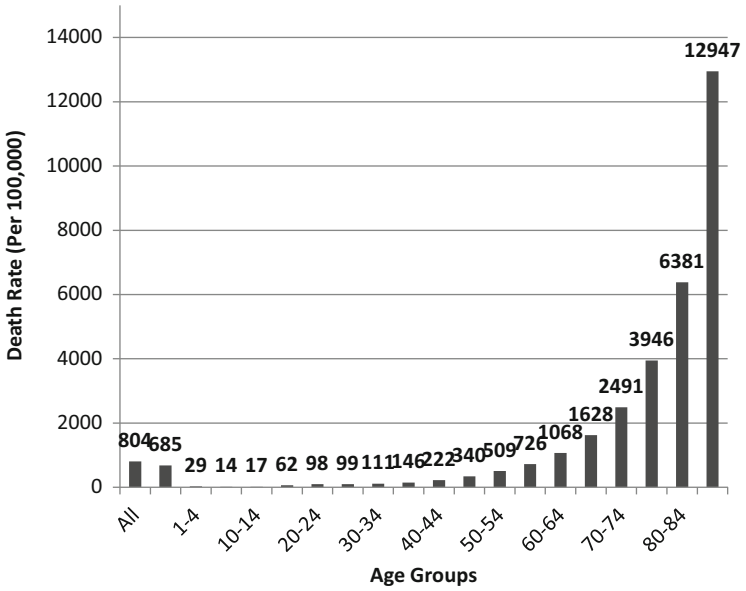

Fig. 3.1 Estimated death rate per 100,000 individuals in 2007 by Age Group. WISQARS data (CDC, 2010a)

(aged 10–14) are marked by particularly low mortality rates with slight increases from middle adolescence to the beginning of midlife (i.e., aged 15–44).

Across all ages, the five leading causes of death in 2007 were heart disease, cancer, stroke, chronic lower respiratory illness (i.e., influenza and pneumonia), and unintentional injuries (CDC, 2010b). Heron et al. (2009) point out that official tabulations of leading causes of death were first published in 1958 and reflect only deaths from identified causes. Over the past 50–60 years of tracking, mortality rates have declined for heart disease, stroke, and to a lesser extent unintentional injuries; mortality from cancer while changing for specific cancers and during particular periods has not demonstrated substantial change over the tracking period.

Thus, if the focus of improving population health was solely on mortality, individuals aged 1–14 or those in the 15–24 age group would hardly be the targets of health initiatives, based on purely cross-sectional comparisons of the lifespan. However, the argument is commonly made that more than any other time in the lifespan, current mortality in adolescence is associated with behavior. Moreover, subsequent mortality in adulthood is often viewed as the outcome of behavioral choices of adolescents and young adults. These conclusions are based on how the causes of mortality vary by age group. For example, the leading causes of death in the population as a whole are heart disease and cancer accounting for about 25 and 23 % of all deaths (CDC, 2010a); rates of mortality from these causes increase sharply with age and are particularly low between ages 10 and 24. In contrast, the top three leading causes of death shift with age with the transition to adolescence.

Specifically, unintentional injuries are the leading cause of death across childhood and adolescence (and into midlife). Childhood cancers are the second most common cause of death from ages 5 to 14. However, homicide shifts from the fourth leading cause of death among 5- to 9-year-olds to the third leading cause of death among 10- to 14-year-olds and subsequently to the second leading cause of death among 15- to 24-year-olds. A similar, but age-delayed, shift occurs for suicide. Prevention of accidents, homicide, and suicide is often quite feasible in the immediate proximity of the event or in the broader context of a child's or an adolescent's life.

In surveys of adolescent health, the broad domains of health-related behaviors tracked are unintentional injury, violence, substance use, reproductive health, nutrition, and physical activity. However, these behaviors are often not distinct. For example, in the category of unintentional injury, one in three deaths among teens is from motor vehicle crashes. Teens are more likely to engage in riskier driving behaviors (e.g., speeding, tailgating), and risk increases dramatically when other teens are passengers in the vehicle. Although these risk factors may be particularly salient to teen driving, it should also be noted that alcohol use was involved in 25 % of accidents in which a teen driver died, and those same teen drivers rarely were wearing seat belts (CDC, 2011). The National Institute on Alcohol Abuse and Alcoholism (NIAAA) found that overall, teen drivers have fatal alcohol-involved accidents twice as often as older drivers (Yi, Chen, & Williams, 2006). In the case of violence, firearms were involved in 84 % of homicides and 46 % of suicides among children and adolescents (CDC, 2010b). In part, this is due to the increased lethality of firearms in comparison with other violent behaviors. However, alcohol is also implicated in substantial numbers of homicides and suicides each year as well as in high-risk sexual behaviors (NIAAA, 2006). Notably, violent behaviors are often examined within the field of adolescent mental health; clearly suicide stems from depression, stress, and substance use problems, but other violent behaviors have been part of what has been termed externalizing problems—a large group of behaviors indicative of poorer adjustment and difficulties in social interactions. Thus, by middle-to-late adolescence, many teens have already engaged in serious health-compromising behaviors as they start driving and experimenting with substance use. In contrast, the majority of younger teens have yet to enter these health trajectories.

In addition, other leading causes of mortality like heart disease and cancer have roots in the transition into adolescence. Heart disease and cancer are often the result of genetic factors that interact with lifestyle, most often behavioral patterns that have been stable over much of adult life. For example, behaviors like tobacco use as well as longer term patterns of diet and exercise are critical components of many adult diseases. Note that nearly all individuals who smoke cigarettes as adults began smoking as adolescents (Johnston, O'Malley, Bachman, & Schulenberg, 2009). As young adolescents increasingly choose their own activities based on interests, peers, and available options, these choices impact current health as well as the formation of longer term behaviors and lifestyles, as in the case of tobacco use. Hence, given that the transition into adolescence is a time of increased decision

making about one's behaviors, the argument is made that lifestyles exhibited in adulthood often have their beginnings during this period with the emergence of greater control over factors associated with injury and subsequent disease such as overweight and obesity.

One challenge in studying the emergence of more autonomous behaviors among young adolescents is determining which behaviors may be temporary or part of the current developmental context and which will have lasting effects. For example, during adolescence, nutritional behaviors may change as youth eat meals or snacks away from home and engage in independent selection of what they eat. Some nutritional choices may have temporary effects with individuals making corrections or better choices at other points in their lives. That is, some young teens may live on a diet of pizza and caffeinated beverages when hanging out with peers but may go on to develop healthy eating patterns in adulthood.

However, some nutritional choices made during adolescence may have lasting impacts on health. For example, in the 2009 Youth Risk Behavior Surveillance reports, only 14.5 % of 9th- to 12th-grade students reported drinking the recommended number of glasses of milk per day (CDC, 2010c); the rate for this behavior has, unfortunately, been low since 1999 when the CDC first began tracking daily consumption of milk among high school students. Related to this, young adolescents typically consume below the recommended levels of calcium often in part because they stop drinking milk and instead drink carbonated beverages that contain phosphoric acid. Not only is milk, a calcium source, removed from the daily diet but also phosphoric acid has been shown to inhibit absorption of calcium from food sources (Wyshak & Frisch, 1994). Wyshak and colleagues demonstrated that consumption of carbonated beverages was associated with bone fractures in adolescent girls and young women (after controlling for several related factors such as age, weight, height, activity level, and total caloric intake). Bone growth increases dramatically during puberty and continues into young adulthood for both males and females. Hence, these behavioral changes commensurate with pubertal development not only have immediate impact on bone health (e.g., increases in bone fractures) but also likely have lasting effects on bone development and health such as propensity for developing osteoporosis in later adulthood.

Of course, some health-promoting behaviors are enhanced by physical growth and increased abilities (e.g., increased stamina and lean muscle mass) that result from puberty. Rates of physical activity often peak in childhood although engagement in sports and physical education in schools helps some adolescents maintain prior levels. For example, 63.8 % of boys and 52.3 % of girls in high school reported that they had been on a sports team in 2009 (CDC, 2010c). About the same percentage of high school students indicated that they had at least one physical education (PE) class per week although the majority of youth do not have PE at school every day (i.e., only 33.3 % have daily PE class). Data for middle school students are similar but demonstrate quite a bit of variation by state for things like PE classes. Given that adolescents are often more likely to have the opportunity to engage in physical activity than adults are, it is concerning that rates

of physical activity typically decline from ages 12 to 21 with additional declines as young adults leave educational environments and enter the workforce (Caspersen, Pereira, & Curran, 2000). In this case, the challenge has been to encourage individuals to continue to be physically active into adulthood. However, it should be noted that even though engagement in some physical activity is occurring for many adolescents, only 24.8 % of boys and 11.4 % of girls in high school report that they met the recommended levels of physical activity in the past week (CDC, 2010c).

In summary, adolescence is a period of low mortality and morbidity for most individuals; most importantly, as children make the transition into adolescence, the leading causes of mortality are preventable. Moreover, mortality and cause of death data provide compelling evidence for the role of behaviors initiated during adolescence in population health, identifying this period as important for preventive intervention for a range of health-related behaviors.

Adolescent Development as the Context for Changes in Health Behaviors

Although national health goals like Healthy People 2020 (Office of Disease Prevention and Health Promotion, 2009) have expanded objectives for those aged 10–24 to include educational achievement, connection to caregivers, and increasing the percentage of vulnerable youth who are prepared to make successful adult transitions, overall these population-based approaches to health lack a strong developmental focus. As we have noted, the transition into adolescence is a period of substantial changes in every, or nearly every, facet of life as the individual transitions from child to adult. In particular, we and others have asserted that the navigation of this transition sets youth onto health trajectories into adulthood (e.g., Graber & Brooks-Gunn, 1996). Drawing on Elder's seminal work on transitions and trajectories (Elder, 1998), transition periods are those in which significant changes in form and function occur. The level of development that has occurred before the transition, the timing of the transition, and the context of the transition interact to set individuals onto behavioral or life trajectories. Although these are not immutable paths, often they require changes in context or subsequent transitions to alter trajectories.

Thus, the impact of transitions can be understood by examining characteristics of the individual and the interaction with context. Drawing on diathesis–stress models of psychopathology, individual factors that predict unhealthy trajectories are often referred to as vulnerabilities; these are predispositional, usually intraindividual, factors such as emotion regulation, physiological responses to stress, or temperament that are in part shaped by experience but often become stable during childhood and adolescence (Ingram & Luxton, 2005). Although this model has most often been applied to adjustment problems, individuals may

develop vulnerabilities that are linked to a range of health-compromising behaviors. In contrast to vulnerabilities that have a causal link to behavior, risk factors are those associated with increased probability of a particular problem (Ingram & Luxton, 2005). For example, gender is a risk factor for depressive problems during adolescence because gender is associated with higher probability of developing these problems but does not explain why the problems occur more often for girls. Protective factors buffer the impact of risk factors and potentially impede the development of unhealthy behavior trajectories. The diathesis–stress model of psychopathology predicts that major transitions or stressful events interact with prior vulnerabilities, resulting in increased problems or poor outcomes in the face of these stressors. In testing this model, the focus is typically on the identification and development of prior vulnerabilities for a particular problem or health-compromising behavior (e.g., poor emotion or behavioral regulation skills, genetic markers).

Whereas diathesis–stress models call attention to the individual factors that set children on the course for healthy and unhealthy transitions, ecological systems theory has been particularly useful as a model for how the contexts of an individual's life shape development in childhood and across the life course (Bronfenbrenner & Morris, 1998; Magnusson & Stattin, 1998). The contexts of adolescents' lives range from the most proximal, such as family and friends, to more distal such as school, community, and society as a whole. Considering not only individual factors in development but also individual differences in contexts of adolescents' lives provides insights into who is at risk for poorer health outcomes and why. Moreover, it is important to identify the contexts that promote healthy trajectories to determine the goals of health promotion as well as to understand why some individuals have better versus poorer health.

In addition, child and adolescent health does not fit into a dichotomy of healthy versus unhealthy, as health is neither a static nor a singular outcome; that is, adolescents may engage in some behaviors that promote healthy outcomes (e.g., use of safety belts, no engagement in violent behavior) but fail to demonstrate this success in other domains (e.g., initiate alcohol use at a young age). In Problem Behavior Theory, Jessor (1992) notes that there is often substantial co-occurrence of several problem behaviors like substance use, delinquency, aggression, and sexual risk behaviors. Given this, it is often feasible to identify common risk and protective factors for these behavioral profiles and hence develop health promotion strategies that impact multiple behaviors. However, studies have demonstrated that behaviors do not always group consistently for different subgroups of adolescents. For example, among urban African American girls, sexual risk behaviors appear to be independent of other problem behaviors (Stanton et al., 1993). Moreover, although problem behaviors are frequently a core focus of adolescent health concerns, Problem Behavior Theory is not a global theory of adolescent health. To date, few studies have examined how other health behaviors like nutrition and physical activity correlate with problem behaviors to determine how the theory might be extended to health more broadly. Thus, for most health behaviors or behavioral groups, the individual and contextual factors that influence each

behavior may be quite different and may change over the course of adolescent development.

Benchmarks of Development from Childhood to Adolescence

Perhaps one of the most substantial challenges faced by adolescents and the adults responsible for promoting their health is determining what the specific developmental goals are for different periods of adolescence and how to determine whether an adolescent has achieved that goal or benchmark for development. Adolescence is often defined by its endpoint: becoming an adult—an endpoint that itself has been defined by different standards depending on expectations for adult roles. This chapter focuses on the entry into adolescence and the early adolescent years. The benchmark or the standard for development in early adolescence may best be thought of as the absence of any single standard. At earlier times in development there seem to be clearer goals like school readiness, a construct that has assessment tools to determine whether cognitive, emotional regulation, and social skills deemed requisite for beginning formal schooling are in place. In contrast, entry into early adolescence might be defined by entry into middle school—a transition that may happen at 5th, 6th, or 7th grade or not at any grade (in K-8 or K-12 schools). The end of early adolescence might also be thought of as the entry into high school, another transition with few clear behavioral indicators of readiness beyond having passed the requisite material in 8th grade.

As we have noted, development in this period is characterized by puberty, a process that will continue over the next 4–5 years before full sexual maturity has been reached. That said, many normatively developing adolescents will continue to experience pubertal development after the early adolescent years. In addition, the brain undergoes significant changes over the course of adolescence, particularly in key areas for more advanced cognitive abilities such as abstract thinking and planfulness. While these developments continue into late adolescence or early adulthood, cognitive changes in early adolescence are typified by emergence of abstract thinking abilities such as increased thinking about possibilities in both a longer future period and other possibilities in the here and now, being able to think about multiple dimensions of an issue rather than a single aspect, and increased time spent thinking about thinking. Piaget and his colleagues documented the emergence of abstract logical reasoning in early adolescence, considering this the distinguishing feature of cognitive differences between childhood and adolescence (Piaget, 1972). However, attainment of these skills is variable and context dependent and may never be achieved by adults who seem to function effectively in the world (Keating, 2004). Hence, clear cognitive standards differentiating an early adolescent from an older adolescent or an adult are not easily defined.

Notably, advancements in abstract thinking allow adolescents to develop greater intimacy in relationships, both friend and romantic, as they apply their cognitive skills to understanding social situations and the perspective of others. Also as

indicated, adolescents become more autonomous behaviorally and emotionally. Early adolescence is typified by increased autonomy from parents and family as youth spend more time with peers (Larson & Verma, 1999). At the same time, they often demonstrate less autonomy with peers, placing greater importance on conformity to peer norms in early adolescence (Costanzo & Shaw, 1966). By late adolescence, individuals can make decisions and have emotional self-reliance that balances one's ideas with input from significant others including family members, friends, and romantic partners.

Thus, given that each aspect of development in early adolescence is a process influenced by individual and contextual factors, the next section considers development of key person-in-context aspects of adolescent development. Moreover, we describe pathways to healthy and less healthy outcomes for young adolescents on the course to attaining the skills needed both later in adolescence and ultimately in adulthood.

Key Person-in-Context Factors

A review of all individual and contextual factors that may influence health during the transition into adolescence is beyond the scope of this chapter. We have selected a few that have been particularly useful and of interest in understanding the health behaviors that have been the focus of attention in the field. In particular, pubertal development influences the course of adolescent physical development as well as a range of behaviors not directly related to the physical changes of puberty but more so to how puberty is experienced in context. Development of social-cognitive decision making skills has also been a central focus of adolescent health research (e.g., Fischhoff, 2008). As research on brain development has advanced, insights into the neurological underpinning of decision making have emerged. In terms of contexts, we will examine those that have been studied most extensively and are most proximal to the individual: family, peers, and emergence of romantic interests.

Puberty. As indicated, puberty is often associated with the transition from late childhood into adolescence and development through early-to-middle adolescence. At the most basic level of analysis, pubertal development results in attainment of sexual maturation, reproductive capacity, and most noticeably a more adult-like form and appearance. However, puberty encompasses internal and external changes in (a) maturation of the reproductive organs and the development of secondary sexual characteristics, (b) the growth spurt, (c) the composition and distribution of fat and muscle tissue in the body, (d) the circulatory and respiratory system associated with increased strength and endurance, and (e) areas of the nervous system and endocrine system associated with reproduction and physical growth (Marshall & Tanner, 1986). The impact of changes in each area on health is easy to see. As noted, the physical and physiological changes of puberty are expected to result in a body that has the potential to be at the peak of physical health by middle-to-late adolescence or early adulthood—strong bones, muscle mass, increased

endurance, and so on. However, group differences in physical capacities, especially by gender, mean that some youth will have bodies better suited to particular health behaviors. For example, boys typically gain more lean muscle mass in comparison with fat than do girls, and effects of testosterone on the circulatory and respiratory systems (specifically hemoglobin in the blood) typically result in a gender advantage for males in physical performance and capacity (Marshall & Tanner, 1986).

Of course, this new and improved body does not form overnight. The development of secondary sexual characteristics typically takes 4–5 years from onset to full maturation. Moreover, as Tanner (1962) observed in his pediatric studies of pubertal development, there is enormous individual variation in timing of onset and rate of pubertal development resulting in wide variation in the level of development within any cohort of children. Variation in pubertal timing, in particular maturing earlier or later than one's peers, has emerged as an important factor in a range of health outcomes and behaviors including sexual behaviors, internalizing and externalizing problems and disorders, and substance use and abuse (for reviews, see Graber, 2003; Mendle, Turkheimer, & Emery, 2007). Mechanisms underlying these effects are related to the experience of off-time development in different contexts such as peer groups or families; mechanisms are also, in some cases, genetic or physiological.

The process of pubertal development is by definition tied to sex and reproduction. However, capacity for reproduction does not inherently mean that someone begins to engage in sexual behaviors. Rather, hormonal changes of puberty may increase arousal (e.g., Udry & Campbell, 1994), and the physical changes of the body may signal to oneself and others that the individual is becoming a sexual being (Graber & Brooks-Gunn, 2002). Yet societal forces and internal regulation often buffer direct effects of puberty on sexual behaviors especially in early adolescence (e.g., Udry & Campbell, 1994). In general, research has found positive correlations between pubertal onset and age of first intercourse. However, the findings are not robust due to minimal significance and nonuniformity of measurements (for a review, see Graber & Brooks-Gunn, 2002).

In contrast, timing of puberty has been consistently associated with the emergence of a range of health-related problems beginning in early adolescence. Whereas early and late maturers of both genders experience periods of elevated subclinical problems during adolescence, in a large longitudinal study, we have found that early-maturing females and late-maturing males have higher rates of not only lifetime history of disorders but also new onset of disorder (Graber, Seeley, Brooks-Gunn, & Lewinsohn, 2004). Specifically, early maturation in girls is associated with both internalizing and externalizing disorders, most consistently with behavioral disorders, anxiety, and depression (for a recent review, see Graber, Nichols, & Brooks-Gunn, 2010). Recent studies have also found that early maturation in girls is linked to greater alcohol, tobacco, and/or substance use (for a review, see Graber, 2003); earlier initiation of and possibly faster progression from cigarette and alcohol use to other drugs (Lanza & Collins, 2002); and higher rates of substance-use disorder by middle adolescence (Graber, Lewinsohn, Seeley, & Brooks-Gunn, 1997). Magnusson, Stattin, and Allen (1985) demonstrated that

association with older peers was an important modifier of links between early maturation in girls and engagement in deviant behaviors. Subsequent studies have found that girls who matured earlier than female peers and associated with older boys as romantic partners had elevated risk for younger ages of onset of sexual intercourse as well as other health-compromising behaviors like substance use (Halpern, Kaestle, & Hallfors, 2007).

It should be noted that variation in timing of puberty itself is influenced by behaviors (e.g., eating and physical activity) as well as psychosocial experiences (Graber et al., 2010), especially for girls. For example, prepubertal sexual abuse actually influences the timing of the onset of puberty in girls. In particular, girls who are sexually abused in childhood before puberty are more likely to mature earlier than non-abused girls (Trickett & Putnam, 1993; Zabin, Emerson, & Rowland, 2005). Onset of sexual behaviors and risk for these early maturers may be more closely tied to the psychological impact of the history of abuse rather than timing of puberty per se.

Early maturation has also been linked to health-compromising behavior among males. Effects of early maturation on psychopathology have been less severe for boys than girls, with early-maturing males having elevated subclinical symptoms rather than higher rates of disorder. However, early maturation in boys has also been associated with increased alcohol use and/or abuse (Wichstrom, 2001; Williams & Dunlop, 1999) as well as delinquent or externalizing behaviors (Ge, Conger, & Elder, 2001; Williams & Dunlop, 1999) in the early adolescent years. In a large epidemiological study, we have found that in young adulthood, early maturing males had significantly higher rates of daily cigarette use than other young men and rates that were higher than rates reported by the CDC for this age group (Graber et al., 2004). In the case of cigarettes, looking older than peers may have made it easier to purchase cigarettes and smoke regularly in public; hence, early-maturing males may have experienced fewer obstacles to smoking and become regular smokers by young adulthood.

Interestingly, as indicated, late maturation appears to confer more serious risk for males during the transition to adulthood. For example, late maturation for boys also seems to confer risk for increased alcohol use and/or abuse in young adulthood (Andersson & Magnusson, 1990; Graber et al., 2004). Notably in a large longitudinal study, late-maturing males had significantly *lower* rates of initiation of substance use in the high school years (Graber et al., 1997), but by age 24, late-maturing males had *higher* rates of substance-use disorders in comparison to on-time maturers. Although these effects emerge after adolescence, there are some indications that late-maturing males experience problems with regulation during the early adolescent period. Recent studies have found that when late-maturing males also experience stress with peers (e.g., rejection, being teased), they have elevated psychopathology; in one study, higher rates of depressive symptoms were reported (Conley & Rudolf, 2009), and in another study, higher rates of aggression were reported for these males (Sontag, Graber, & Clemans, 2011). How such experiences in the early adolescent period may translate over time to substance-use problems in young adulthood is not clear, but these studies suggest

that emotional responses to peer stress are not well regulated by late-maturing males.

Brain development and decision making. Along with the biopsychosocial changes of puberty, research on neuro-cognitive development during adolescence has shed light on emotion and cognitive control functions associated with decision making about health-related behaviors and vulnerability to health problems. New technologies have made it possible to identify both structural and functional brain development occurring during adolescence that is related to changes in decision making more generally (for a review, see Paus, 2009) as well as processes that relate directly to health outcomes.

Although postnatal brain development is considered to occur primarily during childhood, maturation of higher level brain regions continues through adolescence and into adulthood. The outer layer of the brain, or the cortex, is made up of what is called gray matter, which consists of nerve cell bodies and dendrites. Many of the subcortical brain structures, such as the basal ganglia and amygdala, are also made up primarily of nerve bodies and therefore appear darker as well. White matter, on the other hand, appears lighter due to the fatty myelin sheath that covers the axons that transfer neuronal signals to neighboring neurons. Gray matter increases rapidly throughout childhood, reaches its highest levels around the onset of puberty, and then begins to decrease. In contrast, white matter shows steady increases through adolescence and into adulthood. The decrease in gray matter and increase in white matter during adolescence suggest that the brain is becoming more specialized and is developing more numerous and more complex neural connections.

A critical aspect of cognitive development emerging during early adolescence is the ability to engage in more complex and abstract problem solving. Perhaps one of the most frustrating aspects of adolescent behavior, to parents and professionals, is that adolescents often demonstrate advanced cognitive skills and the ability to reason while making decisions that put their health at risk. A wealth of social-cognitive studies has identified cognitive, contextual, or social factors that lead to faulty decision making by young adolescents (e.g., Fischhoff, 2008). Recent studies of brain development have shown that areas of the brain attain "adult" levels of development at different times in adolescence and young adulthood, much like in puberty. The adult brain, however, continues to change across the lifespan both developmentally and in response to experiences.

In terms of adolescent decision making, development of the prefrontal cortex (PFC) throughout adolescence has been hypothesized as most salient to the ability to be goal directed and to engage in impulse control (Casey, Getz, & Galvan, 2008; Paus, 2009). Both capacities are associated with better decision making about health-compromising behaviors (Fischhoff, 2008). In addition, Casey et al. (2008) have recently demonstrated substantial changes in maturation of the dopaminergic system, particularly during early adolescence. Dopaminergic systems are associated with processing of social information; these neural developments are linked to increased sensation seeking and reward sensitivity during early to middle adolescence (e.g., Spear & Varlinskaya, 2010; Steinberg, 2008). Whereas these social-emotional aspects of brain development occur in early adolescence, neural

development in the PFC, and associated cognitive control processes, progresses across the adolescent decade. Casey et al. (2008) hypothesize that this developmental lag may result in heightened responses to socio-emotional cues with lagging cognitive controls. Such research advances have provided potential explanations for why young adolescents, who can demonstrate good skills in hypothetical decision making situations, may still have difficulty regulating risk-taking and impulsive behavior in real-world situations.

Some neural changes may be specific to particular health risks such as alcohol consumption or other drug use. Groundbreaking work by Spear (2011), based on a rat model of adolescence, indicates that two processes are at work that heighten the risk for unhealthy alcohol consumption during adolescence. One is that adolescents seem to be less sensitive to the aversive effects of alcohol, such as motor impairment or hangovers; that is, it may require higher rates of consumption to reach the same level of aversive effects as would be seen in adults. In fact, Spear suggests that adolescents are less likely to pass out, making them more susceptible to alcohol poisoning when drinking large amounts of alcohol in brief periods. The other factor that puts adolescents at unique risk is that they have increased sensitivity to the reward effects of alcohol. Increased social facilitation is one reward that appears to be enhanced in the adolescent brain; in recent studies, adolescent rats demonstrated greater effects of positive reward in comparison with adult rats (Spear & Varlinskaya, 2010). Moreover, these processes may be heightened by the experience of stress, with adolescents feeling greater anxiety reduction via alcohol consumption.

Although these processes seem to occur as part of normative development, there are also indications that individual differences in related processes like sensation seeking are salient to substance use as well as externalizing behaviors. Notably, animal models and recent studies (e.g., Steinberg et al., 2008) suggest that sensation seeking increases during adolescence at the mean level for the population of adolescents. In our own work (Lynne-Landsman, Graber, Nichols, & Botvin, 2011), we have found support for three patterns of development of this construct during adolescence. In a study of urban, minority adolescents, sensation seeking (tapped by inhibitory control and thrill seeking) increased from 6th to 8th grade for about 60 % of the sample; in contrast about 20 % showed stable, low levels of sensation seeking and the third trajectory had the highest levels at 6th grade and maintained these levels over time. These trajectories were predictive of increases in aggression, delinquency, and substance use, in the expected directions over time: that is, those in the high stable group had the greatest increases in the three risk behaviors, those in the "increaser" group had increasing rates of the three risk behaviors (although never as high as the high stable group), and those in the low stable group had low rates of the three risk behaviors.

As demonstrated in the individual differences in sensation seeking at the entry into adolescence (Lynne-Landsman et al., 2011), it is also the case that all young adolescents do not have equivalent neurological processing at the entry into adolescence. Those who enter adolescence with heightened risk-taking or thrill-seeking tendencies may be particularly at risk for poorer decision making during

adolescence. Hence, individual differences in brain development and commensurate cognitive and emotional processing changes during adolescence have been identified as possible vulnerabilities for a range of health-related behaviors and problems. Behavioral disorders such as attention deficit hyperactivity disorder and conduct disorders as well as impulse-control disorders, such as obsessive–compulsive disorder, have been linked to abnormalities in PFC activity (Friedlander & Desrocher, 2006; Halperin & Schulz, 2006). Such vulnerabilities may put some individuals at risk for a range of health-related problems during adolescence and beyond.

Family context. As noted, families are the first exposure to behaviors that promote or compromise health. For any of the focal areas of health that we explore, families, most often primary caregivers, model important health behaviors—tobacco use, dietary behaviors, and physical activity—as well as promote health-related behaviors in the home. Promotion of health behaviors also includes restrictions placed on unhealthy behaviors (e.g., limiting physical inactivity like TV viewing or computer use). During the transition into adolescence, youth develop more independence and spend less time with parents and family and more time with peers. However, despite this shift, parents still exert strong influences on adolescent behavior especially in areas like educational and career goals (Laursen & Collins, 2009). Along with shaping values and goals, parents monitor child and adolescent behavior as well as the selection of social environments and interactions that their adolescent children make. Thus, there are several pathways through which families exert influences on the health-related behaviors of young adolescents.

For example, adolescents who grow up in single-parent households or are raised by non-parental family members such as a grandparent have higher rates of conduct and emotional disorders and are more likely to initiate sexual intercourse at a younger age than those living with both parents (Miller, Benson, & Galbraith, 2001). Such effects may be explained by the fact that single parents likely have less time to monitor adolescent behavior and social contexts and may have less close relationships with teens as they are more likely to experience economic stress such as low income (for a more nuanced discussion of family stress, see Conger & Donnellan, 2007). In the case of sexuality, some single parents may be modeling dating and sexual activity, thereby influencing adolescent attitudes (Moore & Chase-Lansdale, 2001; Newcomer & Udry, 1984).

Regardless of family structure, low levels of parental monitoring have been linked to adolescent engagement in many risk behaviors (Dishion & McMahon, 1998; Kerr & Stattin, 2000). As might be expected for sexual behaviors (as well as other health risk behaviors), more parental or adult supervision and regulation are associated with healthier behavioral choices such as continued abstinence or older age of first sexual intercourse as well as lower risk sexual behaviors including use of contraception and having fewer partners over time (Miller et al., 2001). In particular, young adolescents who spend time in unsupervised interactions with other-gender peers are at risk for early initiation of sexual behaviors (O'Sullivan, 2005) as well as early initiation of tobacco and alcohol use (Graber, 2010). However, adolescents' own interests are also an important component of being in supervised

versus unsupervised settings. O'Sullivan and Brooks-Gunn (2005) found that young adolescent girls who reported greater sexual interest were more likely to spend time in unsupervised settings than were other girls regardless of abstinence attitudes. Thus, some adolescents appear to be more motivated to evade supervision. At the same time, excessive or coercive parental control is associated with higher levels of sexual risk behaviors among adolescents (Miller et al., 2001), indicating that more supervision is not always the best answer to helping young adolescents develop healthy sexual behaviors.

Moreover, the quality of parent–child interactions has pervasive effects on adolescent health behaviors. Even though parent–child conflict appears to become more heated during early adolescence and commensurate with pubertal changes, many parents and adolescents maintain close, supportive relationships (e.g., Paikoff & Brooks-Gunn, 1991; Steinberg, 2001). Children who enter adolescence with poorer quality family relationships are likely to continue to have low-warmth and high-conflict relationships and higher risk for health-compromising behaviors during adolescence (Steinberg, 2001). For example, parental conflict and overbearing or highly critical parenting styles are associated with increased prevalence of both internalizing and externalizing disorders (Berg-Nielsen, Vikan, & Dahl, 2002; Marmorstein & Iacono, 2004). For sexual behaviors, most research indicates that parent–child closeness, support, and warmth are associated with teens engaging in healthier behaviors like remaining abstinent, postponing sexual intercourse, having fewer partners, and using contraception—all of which decrease the risk of pregnancy and sexually transmitted diseases (STDs) (DeVore & Ginsberg, 2005; Wight, Williamson, & Henderson, 2006).

Heritability and gene–environment interactions. Aspects of parenting or family effects on behavior may be accounted for by genetic factors in that genes play a role in parenting behaviors and health-related behaviors that parents model. As indicated, in the diathesis–stress model, diathesis, or underlying vulnerabilities, interacts with stressors to predict psychopathology (e.g., Ingram & Luxton, 2005). Studies of gene–environment interactions are shedding light on the developmental process through which vulnerabilities and risks are conferred within the family context. For example, a functional polymorphism in the promoter region of the serotonin transporter (5-HTT) gene may be predictive of higher rates of depression and suicidality in adolescence and young adulthood when this genetic polymorphism occurs in individuals who also experience significant stress in their family environment during childhood (e.g., maltreatment; Caspi et al., 2003). Of course, parental characteristics such as their own temperament or vulnerabilities to environmental stressors influence not only parents' likelihood of developing a depressive problem but also their parenting behaviors that may contribute to stressful family environments for children.

In the area of substance use, it is well established that family history of abuse or dependence is predictive of adolescent problems. Chassin and Belz (2000) and Chassin, Hussong, and Beltran (2009) indicate that family history is associated with the risk for early onset of alcohol use (and other substances), persistence of alcohol disorder, and faster progression from initiation to problems. However, heritability

estimates for substance-use problems vary widely based on sample—clinical versus community—and substance; moreover, heritability estimates have the same constraints of all behavioral genetics studies (for a discussion of constraints, see Eisenberg, 1998). Hence, attention has turned to examining genes that may be related to underlying processes salient to development of substance-use problems as described above in models for vulnerability to depressive problems. In the case of substance abuse, genes linked to reward sensitivity and inhibitory control/regulation have been the focus of much of this work (Chassin et al., 2009). Notably, many family factors predictive of substance use in adolescence are also predictive of aggression in childhood; such behaviors often precede substance initiation developmentally. Hence, these family factors may be indicative of a more general genetic effect on regulatory processes rather than just substance use.

These are just a few of the ways in which genetics may play a role in the transmission of vulnerability from parents to children for the development of health-related problems. As such, they highlight the need to consider the interplay between genes and family environment in future studies of adolescent health.

Peers and intimate partners. As noted, peers play an increasingly important role in the lives of adolescents, and conformity to peer norms seems to peak in the early adolescent period (Berndt, 1979; Steinberg & Monahan, 2007). At this time, small same-sex dyads or cliques are pervasive in peer environments and are one venue for peers to exert strong influence over one another. In addition, adolescents increase their association with larger mixed-gender groups around this time. These developmental experiences have led to a focus on peer pressure as the mechanism through which adolescents embark on health risk behaviors. In fact, nearly all prevention programs targeting sexual risk, substance use, aggression, or other behaviors include programming on resisting peer pressure, that is, refusal skills or some aspect of peer contexts (e.g., Dishion & Loeber, 1985; Kandel & Andrews, 1987). However, Berndt (1996) and others (Berten & Van Rossem, 2011) have demonstrated that peers rarely need to use overtly coercive means to encourage engagement in various activities. Often individuals make behavioral choices to emulate peers or seek peers because of a desire to be part of a group that engages in particular types of behaviors (e.g., partiers). Thus, as with families, peers may help teens develop health-promoting or health-compromising skills. Most studies have focused on the health-compromising influences with numerous studies on the deleterious effects of associating with peers who engage in deviant behaviors (Sussman, Pokhrel, Ashmore, & Brown, 2007).

Across studies, association with deviant peers is a consistent correlate of aggressive and delinquent behaviors (e.g., Elliott, Huizinga, & Ageton, 1985; Moffitt & Caspi, 2001) as well as early onset of sexual activity (French & Dishion, 2003). It has been argued that deviant peers demonstrate for one another tolerance of problem behaviors, and as such being involved with such a group exposes adolescents to a variety of deviant behaviors and provides encouragement to participate in these behaviors. In fact, Elkington, Bauermeister, Brackis-Cott, Dolezal, and Mellins (2009) found more robust findings for peer influence than caregiver influence on sexual behavior. They suggested that adolescents who perceived their peers

to be involved in sexual behaviors were more likely to engage in similar behaviors themselves.

At the same time, adolescent peer groups and close friends show dramatic instability over the young adolescent period as youth explore their identities and find commonalities with others going through the same process of identity exploration (Cairns & Cairns, 1994). As such, adolescent engagement in health-compromising behavior is in part based on an adolescent seeking out youth who already engage in these behaviors or who are perceived as sharing an interest in exploring these behaviors. These personal attributes can affect adolescents' behaviors as much as or more than peer interactions. Although there is a co-occurrence of deviant behaviors and close friend deviant behaviors, peer influence happens via multiple pathways, both proximal and distal, and includes reciprocal influences among groups of friends.

Although peers may be "bad influences," more often peers are valued for the support they provide—someone to hang out with, someone to share feelings with, and so forth. The emerging emotional and social reasoning skills that are developing over adolescence afford shifts to greater intimacy in peer relationships and the potential for these relationships to have greater impact on the emotional well-being of youth (e.g., Berndt, 1996; Deater-Deckard, 2001). And, although intimacy in relationships may be a source of support, it also may be a source of distress especially if relationships are hurtful or not supportive (Deater-Deckard, 2001). Adolescents of both genders who report feeling part of a peer group have lower levels of internalizing and externalizing symptoms (Neuman, Lohman, & Neuman, 2007); high social status in the peer group has also been linked to lower reports of anxiety and depression. Interestingly, young adolescents with low social status did not have greater risk for internalizing problems as long as they reported feeling accepted by peers (La Greca & Harrison, 2005). Thus, while early adolescence is a period when peer groups and status are particularly important, the quality of relationships and peer acceptance seem to be the most salient aspects of the peer environment for well-being and health.

Negative peer relationships, on the other hand, can have a significant impact on an adolescent's emotional well-being. Victimization, lack of peer support, and problems within close friendships have been associated with increased levels of anxiety and depression in both genders, though the relationship is strongest in girls (La Greca & Harrison, 2005; Sontag & Graber, 2010). In addition, boys, but not girls, who reported high levels of physical victimization, had significantly higher levels of delinquent behaviors (Sullivan, Farrell, & Kliewer, 2006).

As teens mature into middle and late adolescence, they begin to place greater importance on romantic relationships and in turn exhibit a shift in time spent with friends to time spent with romantic partners (e.g., Zimmer-Gembeck, 1999). Much like friendships, romantic partners have been identified as conduits to unhealthy behaviors depending on the characteristics of the partner and the quality of the relationship (Collins, Welsh, & Furman, 2009). For example, positive experiences with dating have shown to be a protective factor against anxiety, but poor relationship quality can lead to higher levels of stress compared with non-dating peers

(LaGreca & Harrison, 2005). Separate from unhealthy relationships, typical adolescent romantic relationships pose challenges for adolescents as they are often both intense and short lived (Collins et al., 2009).

Younger adolescents, in particular, may be more overwhelmed by the emotional strain of relationships as they have less experience managing the intense feelings associated with romantic relationships (Collins et al., 2009). For example, most girls have boyfriends who are somewhat older; when partners are more than 2 years older than girls, health risks increase, especially for younger adolescent girls. In particular, having an older boyfriend has been linked to early sexual activity, drinking, drug use, delinquency, and difficulty in school (Young & d'Arcy, 2005). Among middle school students, girls with older boyfriends had more unwanted sexual advances, had more friends who were sexually active, and were more likely to have had intercourse than other girls (Marín, Coyle, Gómez, Carvajal, & Kirby, 2000); these girls were also more likely to be earlier maturers. Often commensurate with romantic relationships, early onset of sexual behaviors has also been associated with increased prevalence of depressive symptoms (Rubin, Gold, & Primack, 2009). Thus, some of the same relational risks that influence sexual health also influence psychological well-being. And engagement in one risk behavior, that is, early sexual onset, impacts other health behaviors.

Summary. This brief overview of important environmental contexts and aspects of psychological development clearly leaves out many other important biopsychosocial experiences of adolescence. School, neighborhood, and media influence the lives of adolescents, their families, and peers. Overall, typical adolescent experience includes a steady stream of new challenges and changes happening in multiple contexts of life. For some youth, these challenges have cumulative effects that may overwhelm coping skills as they try to navigate new relationship demands, increase conflict in family relationships, and potentially heighten emotional sensitivities to social cues. Influences may also be synergistic with multiple contexts promoting common behavioral outcomes as in cases where parents, peers, and teachers or others support physical activity, academic achievement, or other aspects of better health.

Efforts in Early Adolescence

As we have stated, the transition into adolescence is a relatively healthy period of the lifespan. Even so, there is a large emphasis on health and prevention during this time because adolescents make more independent behavioral choices, including those related to health. The increased responsibilities and independent belief formation make the transition into adolescence and the adolescent years an important time for prevention. Although these years are heavily targeted by a variety of prevention programs, our examples are drawn from drug use, pregnancy, and STD prevention programs because those have been the most widely implemented and evaluated.

Prevention programs that target adolescents are often outcome specific and take advantage of the transitional period of adolescence to establish healthy behaviors as well as improve accurate perceptions of health risk and knowledge of normative behaviors. These programs are frequently designed for and implemented in school settings because these are established places of learning and skill development (Kirby, 2007; Soole, Mazerolle, & Rombouts, 2008). However, programs are also implemented, although less frequently, through community organizations such as faith-based organizations, housing projects, and community centers. Just as programs have varied the venue for reaching adolescents, programs have also used many different approaches to delivering program curricula including adolescent-only programs, parent and youth programs, media-based programs such as the Partnership for a Drug-Free America, clinic-based programs, and more general programs such as service learning or positive youth development programs (for a review of positive youth development programs, see Roth & Brooks-Gunn, 2003).

Although the number of prevention programs available for teens is staggering, there are still many obstacles to creating and implementing effective programs, particularly in school settings. First, many programs are created based solely on theory and are never empirically tested. And when programs are empirically tested, the results often indicate that they do not achieve the desired long-term outcomes. For example, pregnancy and STD prevention programs often increase knowledge about STDs and some even delay initiation of sexual intercourse by a few months, but many do not change behaviors (for a review of evaluation studies assessing pregnancy and STD prevention programs, see Kirby, 2007; for evaluation studies of abstinence-only programs, see Mathematica Policy Research Inc., 2007).

Second, to guarantee efficacy, programs must be implemented exactly as tested in the clinical trials. However, the testing conditions are typically far more optimum (e.g., ample funding and time along with high-quality instructors) than in real-world application settings (Schinke & Cole, 2002). Furthermore, many programs need refresher or booster sessions every year to maintain results; schools often find these session undesirable as they can be expensive and time consuming.

Third, factors unrelated to actual efficacy and outcomes sometimes contribute to prevention program selection and implementation. For example, community organizations and parents often have strong opinions and beliefs about which prevention programs should be taught in schools, and thus schools endure a great amount of pressure to use such programs. However, many programs that are popular in the community—such as D.A.R.E. and the majority of empirically tested abstinence-only sex programs—have empirical evidence demonstrating that they do not achieve their desired goals (Clayton, Cattarello, & Johnstone, 1996; Mathematica Policy Research Inc., 2007).

Finally, it is important to note that schools target many problem behaviors at once and often use a variety of methods for each problem behavior, in part because although programs are successful, their effect sizes are often small. However, the research is limited to evaluating the effectiveness of one program implemented in isolation. Thus, we do not know about the additive effects and interactions that result when multiple programs are implemented at one school. More research is

needed to help schools design successful comprehensive packages of prevention programs.

Despite these concerns in prevention programming, there are commonalities across programs that work in improving health outcomes for young adolescents. Most importantly, such programs seemed to be based in the application of developmental concepts, in particular the interplay between individual and contextual factors that shape developmental outcomes.

As these programs are shaping developmental outcomes, the timing of the programs is important in determining the programs' level of success. Research on various types of substance use prevention programs suggests that programs implemented during the middle school years have the greatest and longest effects on substance use-related behaviors (e.g., Gottfredson & Wilson, 2003; Petrie, Bunn, & Byrne, 2007; Soole et al., 2008). However, Porath-Waller, Beasley, and Beirness (2010) found that, for programs dealing exclusively with cannabis use, programs during the high school years are more effective. Perhaps this is because of the later onset of cannabis use when compared with alcohol and tobacco onset, which is targeted in more wide-ranging prevention programs.

The effectiveness of sex education programs also relates closely to the age range of the participants. Contraceptive programs have better outcomes when the participants are older (Franklin, Grant, Corcoran, O'Dell Miller, & Bultman, 1997). Again, older adolescents are more likely than younger adolescents to be sexually active, and thus contraceptive use is more applicable, whereas younger adolescents are less likely to be sexually experienced and can benefit more from a greater emphasis on delaying sexual activity or abstinence (Kirby, 2007). Overall, there may not be one ideal age during adolescence at which to administer these various prevention programs. Instead, the most successful programs tailor their content and goals to the developmental characteristics of the participants.

Beyond identifying developmentally appropriate content and goals, those implementing prevention programs must also select one with the best content model (e.g., affective, informational, social influence models). However, it has been suggested that the most effective models are effective because of the delivery methods of the content, as opposed to the content itself (Kirby, 2007; Wilson, Gottfredson, & Najaka, 2001). As such, we see that successful prevention programs across many areas use a wide range of interactive methods (as opposed to didactic methods) to deliver the program's skills and knowledge to participants (Kirby, 2007; Petrie et al., 2007; Porath-Waller et al., 2010; Soole et al., 2008; Tobler et al., 2000; Wilson et al., 2001). Thus, given that young adolescents are in the process of developing new skills to meet new challenges, successful programs provide the opportunities for actively practicing skills, observing others, and receiving feedback on the strategies they try.

Conclusions

In the course of this chapter, the goal has been to discuss a developmental focus on health during the transition from childhood into adolescence, how models of development inform our understanding of how youth develop healthy and unhealthy attitudes and behaviors, and the key person-in-context factors that influence health in this developmental period. Although we have selected specific person-in-context factors that seem to provide a basis for individual differences in several health outcomes (e.g., sexual activity, substance use), other factors also likely impact health behaviors of adolescence. We note that schools, media, cultural norms, state and federal policies, and societal norms likely have unique influences on the attitudinal and behavioral health practices of adolescents. Also, many of our examples of person and context influences on behavior are limited to sexuality, internalizing and externalizing problems, and substance use because there has been quite a bit of research and prevention efforts around these issues. As noted, substantial attention has focused in recent years on changes in rates of obesity among the US population of adults, children, and adolescents. Notably, most discussions of obesity consider multiple factors that influence this outcome and would likely fit well under Bronfenbrenner's ecological systems theory. For example, historical changes in food production and marketing, community differences in access to different types of foods, opportunities for physical activity, family practices in nutrition and physical activity, and individual opportunities for engaging in sedentary activity like playing video games contribute to changes in weight and obesity (for recent developmentally informed discussions of this topic, see Paxson, Donahue, Orleans, & Grisson, 2006; Philipsen & Brooks-Gunn, 2008). As such, our hope is that rather than reviewing each health outcome, our models will be extended to discussions of other outcomes and shed light on common and unique developmental processes underlying multiple health outcomes.

Although we have argued that the entry into adolescence is a time when health behaviors are especially important because of their potentially longer term impact, examination of prior development in early and middle childhood to understand factors that influence the trajectory for health is essential. Several of the health behaviors we have discussed are salient across the lifespan and exhibit much continuity, impacting later development in adulthood as well as being predicted from earlier childhood experiences. The present discussion was limited to a single period of life. While the transition into adolescence is a time when individuals begin to exert more autonomy in their health-related behaviors and make more independent choices, preexisting behavioral patterns are an important factor affecting health behaviors of adolescents. Notably, socializing experiences and genetic factors interact in the development of many health behaviors, and both contribute to the establishment of behavioral patterns before and during adolescence.

Across the correlates and possible predictors of the health outcomes, it is clear that recent studies have been more integrative and have examined multiple factors simultaneously. Such approaches have made it feasible to compare relative

influences of predictors on different health outcome as well as influences common to several outcomes. However, no single study can examine in depth each factor discussed in this review, but focused investigations that study pieces of different models and make comparative tests of models will continue to advance this area of research. Although several new findings have emerged in recent years and certainly since the original studies of adolescent health were begun, there are still numerous gaps in the literature, especially in terms of integration of constructs.

Moreover, the critical factor has been whether healthy behaviors can be promoted effectively and whether unhealthy behaviors can be prevented. As indicated, numerous prevention/health promotion programs have been developed and implemented among young adolescents. However, nearly all of these programs are domain specific (i.e., sexual risk versus violence or drug use). For some health outcomes, specific targets of intervention are domain specific, but in many others the potential targets, such as counteracting reward sensitivities, parental monitoring, and quality of parent–adolescent relationships, seem to be fundamental pathways to several health outcomes. To date, because of a range of cost, time, and related constraints, health promotion initiatives have yet to capitalize on the synergies across domains. At the same time, long-term cost benefits of establishing healthy "lifestyles" during adolescence support the need for investment in such initiatives.

References

Andersson, T., & Magnusson, D. (1990). Biological maturation in adolescence and the development of drinking habits and alcohol abuse among young males: A prospective longitudinal study. *Journal of Youth and Adolescence, 19*, 33–41. doi:10.1007/BF01539443.

Berg-Nielsen, T. S., Vikan, A., & Dahl, A. A. (2002). Parenting related to child and parental psychopathology: A descriptive review of the literature. *Clinical Child Psychology and Psychiatry, 7*, 529–552. doi:10.1177/1359104502007004006.

Berndt, T. J. (1996). Transitions in friendship and friends' influence. In J. A. Graber, J. Brooks-Gunn, & A. C. Petersen (Eds.), *Transitions through adolescence: Interpersonal domains and context* (pp. 57–84). Hilllsdale, NJ: Lawrence Erlbaum Associates.

Berndt, T. J. (1979). Developmental changes in conformity to peers and parents. *Developmental Psychology, 15*, 608–616. doi:10.1037/0012-1649.15.6.608.

Berten, H., & Van Rossem, R. (2011). Mechanisms of peer influence among adolescents: Cohesion versus structural equivalence. *Sociological Perspectives, 54*, 183–204. doi:10.1525/sop.2011.54.2.183.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (5th ed., Vol. 1, pp. 685–759). Hoboken, NJ: Wiley.

Cairns, R. B., & Cairns, B. D. (1994). *Lifelines and risks: Pathways of youth in our time*. New York: Cambridge University Press.

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review, 28*, 62–77. doi:10.1162/jocn.1997.9.6.835.

Caspersen, C. J., Pereira, M. A., & Curran, K. M. (2000). Changes in physical activity patterns in the United States, by sex and cross-sectional age. *Medicine and Science in Sports and Exercise, 32*(9), 1601–1609.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H. L., et al. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science, 301*, 386–389. doi:10.1126/science.1083968.

Centers for Disease Control and Prevention (2010a). Web-based injury statistics query and reporting system (WISQARS). Retrieved from http://www.cdc.gov/ncipc/wisqars.

Centers for Disease Control and Prevention. (2010b). *10 Leading causes of death by age group, United States—2009*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Centers for Disease Control and Prevention. (2010c). Youth risk behavior surveillance—United States, 2009. *Morbidity and Mortality Weekly Report, 59*((SS-5)), 1–36.

Centers for Disease Control and Prevention. (2011). *Teen drivers fact sheet*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Chassin, L., & Belz, A. (2000). Substance use and abuse outcomes in children of alcoholics. In H. E. Fitzgerald, B. M. Lester, & B. S. Zuckerman (Eds.), *Children of addiction: Research, health, and public policy issues* (pp. 193–216). New York: RoutledgeFalmer.

Chassin, L., Hussong, A., & Beltran, I. (2009). Adolescent substance use. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 723–763). New York: Wiley.

Clayton, R. R., Cattarello, A. M., & Johnstone, B. M. (1996). The effectiveness of Drug Abuse Resistance Education (Project DARE): 5-year follow-up results. *Preventive Medicine, 25*, 307–318. doi:10.1006/pmed.1996.0061.

Collins, W. A., Welsh, D. P., & Furman, W. (2009). Adolescent romantic relationships. *Annual Review of Psychology, 60*, 631–652. doi:10.1146/annurev.psych.60.110707.163459.

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annual Review of Psychology, 58*, 175–199. doi:10.1146/annurev.psych.58.110405.085551.

Conley, C. S., & Rudolph, K. D. (2009). The emerging sex difference in adolescent depression: Interacting contributions of puberty and peer stress. *Development and Psychopathology, 21*, 593–620. doi:10.1017/S0954579409000327.

Costanzo, P. R., & Shaw, M. E. (1966). Conformity as a function of age level. *Child Development, 37*, 967–975. doi:10.2307/1126618.

Deater-Deckard, K. (2001). Annotation: Recent research examining the role of peer relationships in the development of psychopathology. *Journal of Child Psychology and Psychiatry, 42*, 565–579. doi:10.1111/1469-7610.00753.

DeRose, L., Graber, J. A., & Brooks-Gunn, J. (2010). Psychological effects of precocious and delayed puberty. In R. Nuss & Y. Frank (Eds.), *Cognitive and behavioral abnormalities of pediatric diseases* (pp. 121–127). New York: Oxford University Press.

DeVore, E. R., & Ginsberg, K. R. (2005). The protective effects of good parenting in adolescents. *Current Opinion in Pediatrics, 17*, 460–465. doi:10.1097/01.mop.0000170514.27649.c9.

Dishion, T. J., & Loeber, R. (1985). Adolescent marijuana and alcohol use: The role of parents and peers revisited. *American Journal of Drug and Alcohol Abuse, 11*, 11–25. doi:10.3109/00952998509016846.

Dishion, T. J., & McMahon, R. J. (1998). Parental monitoring and the prevention of child and adolescent problem behavior: A conceptual and empirical formulation. *Clinical Child and Family Psychology Review, 1*, 61–75. doi:10.1023/A:1021800432380.

Eisenberg, L. (1998). Nature, niche, nurture: The role of social experience in transforming genotype into phenotype. *Academic Psychiatry, 22*(4), 213–222. doi:10.1007/BF03340021.

Elder, G. H., Jr. (1998). The life course and human development. In W. Damon (Series Ed.) & R. M. Lerner, (Vol. Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (pp. 939–991). New York: Wiley.

Elkington, K. S., Bauermeister, J. A., Brackis-Cott, E., Dolezal, C., & Mellins, C. A. (2009). Substance use and sexual risk behavior in perinatally immunodeficiency virus-exposed youth: Roles of caregivers, peers, and HIV status. *Journal of Adolescent Health, 45*, 133–141. doi:10.1016/j.jadohealth.2009.01.004.

Elliott, D. S., Huizinga, D., & Ageton, S. (1985). *Explaining delinquency and drug use*. Beverly Hills, CA: Sage Publications.

Fischhoff, B. (2008). Assessing adolescent decision-making competence. *Developmental Review, 28*, 12–28. doi:10.1016/j.dr.2007.08.001.

Franklin, C., Grant, D., Corcoran, J., O'Dell Miller, P., & Bultman, L. (1997). Effectiveness of prevention programs for adolescent pregnancy: A meta-analysis. *Journal of Marriage and the Family, 59*, 551–567. doi:10.2307/353945.

French, D. C., & Dishion, T. J. (2003). Predictors of early initiation of sexual intercourse among high-risk adolescents. *Journal of Early Adolescence, 23*, 295–315. doi:10.1177/0272431603254171.

Friedlander, L., & Desrocher, M. (2006). Neuroimaging studies of obsessive-compulsive disorder in adults and children. *Clinical Psychology Review, 26*, 32–49. doi:10.1016/j.cpr.2005.06.010.

Ge, X., Conger, R. D., & Elder, G. H. (2001). The relation between puberty and psychological distress in adolescent boys. *Journal of Research on Adolescence, 11*, 49–70. doi:10.1111/1532-7795.00003.

Gottfredson, D. C., & Wilson, D. B. (2003). Characteristics of effective school-based substance abuse prevention. *Prevention Science, 4*, 27–38. doi:10.1023/A:1021782710278.

Graber, J. A. (2003). Puberty in context. In C. Hayward (Ed.), *Gender differences at puberty* (pp. 307–325). New York: Cambridge University Press. doi:10.1017/CBO9780511489716.015.

Graber, J. A. (2010). *Evaluator final report*. Heartland Rural Health Network (CBAE Program, Grant No. 90AE0276).

Graber, J. A., & Brooks-Gunn, J. (1996). Transitions and turning points: Navigating the passage from childhood through adolescence. *Developmental Psychology, 32*(4), 768–776. doi:10.1037/0012-1649.32.4.768.

Graber, J. A., & Brooks-Gunn, J. (2002). Adolescent girls' sexual development. In G. M. Wingood & R. J. DiClemente (Eds.), *Handbook of women's sexual and reproductive health* (pp. 21–42). New York: Kluwer Academic/Plenum Publishers.

Graber, J. A., Lewinsohn, P. M., Seeley, J. R., & Brooks-Gunn, J. (1997). Is psychopathology associated with the timing of pubertal development? *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1768–1776. doi:10.1097/00004583-199712000-00026.

Graber, J. A., Nichols, T. R., & Brooks-Gunn, J. (2010). Putting pubertal timing in developmental context: Implications for prevention. *Developmental Psychobiology, 52*, 254–262. doi:10.1002/dev.20438.

Graber, J. A., Seeley, J. R., Brooks-Gunn, J., & Lewinsohn, P. M. (2004). Is pubertal timing associated with psychopathology in young adulthood? *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 718–726. doi:10.1097/01.chi.0000120022.14101.11.

Halperin, J. M., & Schulz, K. P. (2006). Revisiting the role of the prefrontal cortex in the pathophysiology of attention deficit/hyperactivity disorder. *Psychological Bulletin, 132*, 560–581. doi:10.1037/0033-2909.132.4.560.

Halpern, C. T., Kaestle, C. E., & Hallfors, D. D. (2007). Perceived physical maturity, age of romantic partners, and adolescent risk behavior. *Prevention Science, 8*, 1–10. doi:10.1007/s11121-006-0046-1.

Heron, M. P., Hoyert, D. L., Murphy, S. L., Xu, J. Q., Kochanek, K. D., & Tejada-Vera, B. (2009). Deaths: Final data for 2006. *National vital statistics report, 57*(14). Hyattsville, MD: National Center for Health Statistics.

Ingram, R. E., & Luxton, D. D. (2005). Vulnerability-stress models. In B. L. Hankin & J. Abela (Eds.), *Development of psychopathology: Stress—Vulnerability perspectives* (pp. 32–46). New York: Sage Publications.

Jessor, R. (1992). Risk behavior in adolescence: A psychosocial framework for understanding and action. *Developmental Review, 12*, 374–390. doi:10.1016/0273-2297(92)90014-S.

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* (Vol. 2). Bethesda, MD: National Institute on Drug Abuse. College students and adults 19–50 (NIH Publication No. 10-7585).

Kandel, D. B., & Andrews, K. (1987). Processes of adolescent socialization by parents and peers. *Substance Use & Misuse, 22*, 319–342.

Keating, D. P. (2004). Cognitive and brain development. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 45–84). New York: Wiley.

Kerr, M., & Stattin, H. (2000). What parents know, how they know it, and several forms of adolescent adjustment: Further support for a reinterpretation of monitoring. *Developmental Psychology, 36*, 366–380. doi:10.1037/0012-1649.36.3.366.

Kirby, D. (2007). *Emerging answers: Research findings on programs to reduce teen pregnancy and sexually transmitted diseases*. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy.

La Greca, A. M., & Harrison, H. M. (2005). Adolescent peer relationships, friendships, and romantic relationships: Do they predict social anxiety and depression? *Journal of Clinical Child and Adolescent Psychology, 34*, 49–61. doi:10.1207/s15374424jccp3401_5.

Lanza, S. T., & Collins, L. M. (2002). Pubertal timing and the onset of substance use in females during early adolescence. *Prevention Science, 3*, 69–82. doi:10.1023/A:1014675410947.

Larson, R. W., & Verma, S. (1999). How children and adolescents spend time across the world: Work, play, and developmental opportunities. *Psychological Bulletin, 125*, 701–736. doi:10.1037/0033-2909.125.6.701.

Laursen, B., & Collins, W. A. (2009). Parent-child relationships during adolescence. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed., pp. 3–42). New York: Wiley.

Lynne-Landsman, S. D., Graber, J. A., Nichols, T. R., & Botvin, G. J. (2011). Is sensation seeking a stable trait or does it change over time? *Journal of Youth and Adolescence, 40*, 48–58. doi:10.1007/s10964-010-9529-2.

Magnusson, D., & Stattin, H. (1998). Person-context interaction theories. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (5th ed., Vol. 1, pp. 685–759). Hoboken, NJ: Wiley.

Magnusson, D., Stattin, H., & Allen, V. L. (1985). Biological maturation and social development: A longitudinal study of some adjustment processes from mid-adolescence to adulthood. *Journal of Youth and Adolescence, 14*, 267–283. doi:10.1007/BF02089234.

Marin, B. V., Coyle, K. K., Gomez, C. A., Carvajal, S. C., & Kirby, D. B. (2000). Older boyfriends and girlfriends increase risk of sexual initiation in young adolescents. *Journal of Adolescent Health, 27*, 409–418. doi:10.1016/S1054-139X(00)00097-5.

Marmorstein, N. R., & Iacono, W. G. (2004). Major depression and conduct disorder in youth: Associations with parental psychopathology and parent–child conflict. *Journal of Child Psychology and Psychiatry, 45*, 377–386. doi:10.1111/j.1469-7610.2004.00228.x.

Marshall, W. A., & Tanner, J. M. (1986). Puberty. In F. Falkner & J. M. Tanner (Eds.), *Human growth: A comprehensive treatise* (2nd ed., pp. 171–209). New York: Plenum Press.

Mathematica Policy Research Inc (2007). *Impacts of four title V, section 510 abstinence education programs: Final report* (MPR Reference No.: 8549–110). Retrieved from http://www.mathematica-mpr.com/publications/PDFs/impactabstinenceES.pdf.

Mendle, J., Turkheimer, E., & Emery, R. E. (2007). Detrimental psychological outcomes associated with early pubertal timing in adolescent girls. *Developmental Review, 27*, 151–171. doi:10.1016/j.dr.2006.11.001.

Miller, B. C., Benson, B., & Galbraith, K. A. (2001). Family relationships and adolescent pregnancy risk: A research synthesis. *Developmental Review, 21*, 1–38. doi:10.1006/drev.2000.0513.

Millstein, S. G., Petersen, A. C., & Nightingale, E. O. (1993). Adolescent health promotion: Rationale, goals, and objectives. In S. G. Millstein, A. C. Petersen, & E. O. Nightingale (Eds.), *Promoting the health of adolescents: New directions for the twenty-first century* (pp. 3–10). New York: Oxford University Press.

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescent-limited antisocial pathways among males and females. *Development and Psychopathology, 13*, 355–375. doi:10.1017/S0954579401002097.

Moore, M. R., & Chase-Lansdale, P. L. (2001). Sexual intercourse and pregnancy among African American girls in high-poverty neighborhoods: The role of family and perceived community environment. *Journal of Marriage and Family, 63*, 1146–1157. doi:10.1111/j.1741-3737.2001.01146.x.

National Institute on Alcohol Abuse and Alcoholism (2006). Underage drinking: Why do adolescents drink, what are the risks, and how can underage drinking be prevented? *Alcohol Alert, 67*, 1–7.

Neuman, B. M., Lohman, B. J., & Neuman, P. R. (2007). Peer group membership and a sense of belonging: Their relationship to adolescent behavior problems. *Adolescence, 42*, 241–263.

Newcomer, S. F., & Udry, J. R. (1984). Mother's influence on the sexual behavior of their teenage children. *Journal of Marriage and Family, 46*, 477–485.

O'Sullivan, L. F. (2005). The social and relationship contexts and cognitions associated with romantic and sexual experiences of early-adolescent girls. *Sexuality Research & Social Policy, 2*(3), 13–24. doi:10.1525/srsp.2005.2.3.13.

O'Sullivan, L. F., & Brooks-Gunn, J. (2005). The timing of changes in girls' sexual cognitions and behaviors in early adolescence: A prospective, cohort study. *Journal of Adolescent Health, 37*(3), 211–219. doi:10.1016/j.jadohealth.2004.08.019.

Office of Disease Prevention and Health Promotion (2009). Healthy People 2020. Retrieved from http://www.healthypeople.gov/2020/topicsobjectives2020/pdfs/AdolescentHealth.pdf.

Ozer, E. M., & Irwin, C. E. (2009). Adolescent and young adult health: From basic health status to clinical interventions. In R. E. Lerner & R. M. Steinberg (Eds.), *Handbook of adolescent psychology: Individual bases of adolescent development* (5th ed., Vol. 1, pp. 618–641). Hoboken, NJ: Wiley.

Paikoff, R. L., & Brooks-Gunn, J. (1991). Do parent–child relationships change during puberty? *Psychological Bulletin, 110*, 47–56. doi:10.1037/0033-2909.110.1.47.

Paus, T. (2009). Brain development. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 95–115). New York: Wiley.

Paxson, C., Donahue, E., Orleans, C. T., & Grisso, J. A. (2006). Introducing the issue. *The Future of Children, 16*, 3–17.

Petrie, J., Bunn, F., & Byrne, G. (2007). Parenting programmes for preventing tobacco, alcohol, or drugs misuse in children <18: A systematic review. *Health Education Research, 22*, 177–191. doi:10.1093/her/cyl061.

Philipsen, N., & Brooks-Gunn, J. (2008). Overweight and obesity in childhood. In T. P. Gullotta & G. M. Blau (Eds.), *Handbook of childhood behavioral issues: Evidence-based approaches to prevention and treatment* (pp. 125–146). New York: Taylor & Francis Group.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development, 15*, 1–12. doi:10.1159/000271225.

Porath-Waller, A. J., Beasley, E., & Beirness, D. J. (2010). A meta-analytic review of school-based prevention for cannabis use. *Health Education & Behavior, 37*, 709–723. doi:10.1177/1090198110361315.

Roth, J. L., & Brooks-Gunn, J. (2003). Youth development programs and healthy development: A review and next steps. In D. Romer (Ed.), *Reducing adolescent risk: Toward an integrated approach* (pp. 355–365). Thousand Oaks, CA: Sage Publications.

Rubin, A. G., Gold, M. A., & Primack, B. A. (2009). Associations between depressive symptoms and sexual risk behavior in a diverse sample of female adolescents. *Journal of Pediatric and Adolescent Gynecology, 22*, 306–312. doi:10.1016/j.jpag.2008.12.011.

Schinke, S., & Cole, K. (2002). Prevention program implementation. In Z. Sloboda & W. J. Bukoski (Eds.), *Handbook of drug abuse prevention: Theory, science, and practice* (pp. 575–586). New York: Springer Science & Business Media, LLC.. doi:10.1007/0-387-35408-5_29.

Sontag, L. M., & Graber, J. A. (2010). Coping with perceived peer stress: Gender-specific and common pathways to symptoms of psychopathology. *Developmental Psychology, 46*, 1605–1620. doi:10.1037/a0020617.

Sontag, L. M., Graber, J. A., & Clemans, K. H. (2011). The role of peer stress and pubertal timing on symptoms of psychopathology during early adolescence. *Journal of Youth and Adolescence, 40*, 1371–1382. doi:10.1007/s10964-010-9620-8.

Soole, D. W., Mazerolle, L., & Rombouts, S. (2008). School-based drug prevention programs: A review of what works. *Australian and New Zealand Journal of Criminology, 41*, 259–286.

Spear, L. P. (2011). Adolescent neurobehavioral characteristics, alcohol sensitivities, and intake: Setting the stage for alcohol use disorders? *Child Development Perspectives, 5*, 231–238. doi:10.1111/j.1750-8606.2011.00182.x.

Spear, L. P., & Varlinskaya, E. I. (2010). Sensitivity to ethanol and other hedonic stimuli in an animal model of adolescence: Implications for prevention science? *Developmental Psychobiology, 52*, 236–243. doi:10.1002/dev.20457.

Stanton, B., Romer, D., Ricardo, I., Black, M., Feigelman, S., & Galbraith, J. (1993). Early initiation of sex and its lack of association with risk behaviors among adolescent African-Americans. *Pediatrics, 92*, 13–19.

Steinberg, L. (2001). We know some things: Parent–adolescent relationships in retrospect and prospect. *Journal of Research on Adolescence, 11*, 1–19. doi:10.1111/1532-7795.00001.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*, 78–106. doi:10.1016/j.dr.2007.08.002.

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: Evidence for a dual systems model. *Developmental Psychology, 44*, 1764–1778. doi:10.1037/a0012955.

Steinberg, L., & Monahan, K. C. (2007). Age differences in resistance to peer influence. *Developmental Psychology, 43*, 1531–1543. doi:10.1037/0012-1649.43.6.1531.

Sullivan, T. N., Farrell, A. D., & Kliewer, W. (2006). Peer victimization in early adolescence: Association between physical and relational victimization and drug use, aggression, and delinquent behaviors among urban middle school students. *Development and Psychopathology, 18*, 119–137. doi:10.1017/S095457940606007X.

Sussman, S., Pokhrel, P., Ashmore, R. D., & Brown, B. B. (2007). Adolescent peer group identification and characteristics: A review of the literature. *Addictive Behaviors, 32*, 1602–1627. doi:10.1016/j.addbeh.2006.11.018.

Tanner, J. M. (1962). *Growth at adolescence* (2nd ed.). Oxford, England: Blackwell.

Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 metal-analysis. *The Journal of Primary Prevention, 20*, 275–336. doi:10.1023/A:1021314704811.

Trickett, P. K., & Putnam, F. W. (1993). Impact of child sexual abuse on females: Toward a developmental, psychobiological integration. *Psychological Science, 4*(2), 81–87. doi:10.1111/j.1467-9280.1993.tb00465.x.

Udry, J. R., & Campbell, B. C. (1994). Getting started on sexual behavior. In D. John and T. Catherine (Series Ed.) & A. S. Rossi (Vol. Ed.), *Sexuality across the life course. MacArthur Foundation series on mental health and development: Studies on successful midlife development* (pp. 187–207). Chicago: University of Chicago Press.

Wichstrom, L. (2001). The impact of pubertal timing on adolescents' alcohol use. *Journal of Research on Adolescence, 11*, 131–150. doi:10.1111/1532-7795.00007.

Wight, D., Williamson, L., & Henderson, M. (2006). Parental influences on young people's sexual behavior: A longitudinal analysis. *Journal of Adolescence, 29*, 473–494. doi:10.1016/j.adolescence.2005.08.007.

Williams, J. M., & Dunlop, L. C. (1999). Pubertal timing and self-reported delinquency among male adolescents. *Journal of Adolescence, 22*, 157–171. doi:10.1006/jado.1998.0208.

Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2001). School-based prevention of problem behaviors: A meta-analysis. *Journal of Quantitative Criminology, 17*, 247–272. doi:10.1023/A:1011050217296.

World Health Organization (1947). *Official records of the World Health Organization no. 2: Proceedings and final acts of the International Health Conference*. Geneva: World Health Organization

Wyshak, G., & Frisch, R. E. (1994). Carbonated beverages, dietary calcium, the dietary calcium/phosphorous ratio, and bone fractures in girls and boys. *Journal of Adolescent Health, 15*, 210–215. doi:10.1016/1054-139X(94)90506-1.

Yi, H., Chen, C. M., & Williams, G. D. (2006). *Surveillance report #76: Trends in alcohol-related fatal traffic crashes, United States, 1982–2004*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.

Young, A. M., & d'Arcy, H. (2005). Older boyfriends of adolescent girls: The cause or a sign of the problem? *Journal of Adolescent Health, 36*, 410–419. doi:10.1016/j.jadohealth.2004.08.007.

Zabin, L. S., Emerson, M. R., & Rowland, D. L. (2005). Childhood sexual abuse and early menarche: The direction of their relationship and its implications. *Journal of Adolescent Health, 36*, 393–400. doi:10.1016/j.jadohealth.2004.07.013.

Zimmer-Gembeck, M. J. (1999). Stability, change, and individual differences in involvement with friends and romantic partners among adolescent females. *Journal of Youth and Adolescence, 28*, 419–438. doi:10.1023/A:1021612907073.

Chapter 4
Adolescence and Early Adulthood

Jungmeen Kim-Spoon and Julee P. Farley

Introduction

Adolescence can be defined as a transitional period between childhood and adulthood. It begins with sexual maturation (the physical/biological changes related to puberty) and ends with the attainment of adult roles and responsibilities. Thus, adolescence involves transitions in social roles (from that of a child to that of an adult) interposed with a multitude of pubertal changes in body and brain (Dahl, 2004). There is no single event or boundary line that denotes the end of childhood or the beginning of adolescence. Instead, the passage from childhood into and through adolescence comprises a set of transitions that unfold gradually and that touch on many aspects of the individual's development (Steinberg, 2002). These transitions involve challenges and opportunities in physical, cognitive, and emotional development.

Adolescence is marked by rapid increases in physical and mental capabilities. In adolescence, individuals have matured beyond the frailties of childhood but have not yet begun any of the declines of adult aging. Compared with young children, adolescents are physically stronger and faster, can better withstand injury and physical stress, and show maturational improvement in cognitive abilities such as reaction time and reasoning. Therefore, in almost all aspects of human development, adolescence may be seen as a developmental period of strength and resilience. However, researchers point out a striking paradox regarding overall health statistics in adolescence (Dahl, 2004). Ironically, overall morbidity and mortality rates increase 200 % from childhood into late adolescence and early adulthood. Major sources of death and disability in adolescence seem to be related to difficulties in the control of behavior and emotions as expressed in the high rates

J. Kim-Spoon (✉)
Department of Psychology, Virginia Polytechnic Institute and State University,
109 Williams Hall (MC 0436), Blacksburg, VA 24061, USA
e-mail: jungmeen@vt.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_4,

of accidents, suicide, depression, homicide, reckless behaviors and health problems related to alcohol and substance abuse, and risky sexual behaviors (Dahl, 2004).

Risk-taking is defined in the developmental literature as engagement in behaviors that are associated with some probability of undesirable results (Boyer, 2006). The developmental periods of adolescence and early adulthood are characterized by heightened vulnerability to risk-taking behaviors that are acknowledged as threats to adolescent health such as engaging in unsafe sexual activity, experimenting with drugs and alcohol, dangerous driving, and severe delinquent and criminal behaviors (Dahl, 2004). Adolescence is also the time of greatest risk for the initiation of drug use and the development of drinking problems (McCord, 1990). Indeed, recent US data indicate that adolescents ages 12–17 reported that 33.9 % use alcohol, 8.4 % use cigarettes, and 14.6 % use marijuana (Substance Abuse and Mental Health Services Administration [SAMHSA], 2005). Sexual risk behaviors represent substantial—though potentially preventable—risks to mortality, health, and functioning (Blum & Nelson-Mmari, 2004; Williams, Holmbeck, & Greenley, 2002). Particularly in the United States, adolescents and young adults are increasingly at risk for infection with HIV and other sexually transmitted diseases (STDs). Recent US data indicate 50,134 cumulative cases of AIDS among people ages 15–24 from the beginning of the epidemic through 2009. From 2006 through 2009, the rates of annual diagnoses of HIV infection increased from 9.6 to 12.0 per 100,000 for adolescents ages 15–19 and from 28.2 to 36.9 per 100,000 for young adults ages 20–24 (Centers for Disease Control and Prevention, 2010). Emotional and behavioral problems also emerge in adolescence. About 14 % of adolescents experience an episode of major depression by age 17, and a similar proportion show growth in delinquency and violence with many of the affected adolescents showing growth in symptoms across multiple domains of functioning (SAMHSA, 2005). The cost to society of adolescent health-risk behaviors (e.g., greater than $1 billion per year spent in the United States on risky behavior prevention programs) requires research that will enhance the capacity of scientists, practitioners, and policy makers to improve public health outcomes in adolescence and the transition to adulthood (Steinberg, 2004, 2008).

This chapter provides a broad account of what is understood about vulnerability during adolescence and young adulthood with respect to health-risk behaviors. As will be demonstrated, developmental research to date has documented great numbers of correlates and predictors of health-risk behaviors among adolescents and young adults including physical vulnerability (e.g., physical condition and physical environment), cognitive vulnerability (e.g., immaturity in brain development and heightened reward seeking), and emotional vulnerability (e.g., emotion dysregulation and avoidance coping).

Normative Benchmarks of Development During Adolescence and Early Adulthood

Physical Development

A host of physical changes occur during adolescence that are largely due to puberty. A release of hormones marks the beginning of puberty, and often the first visible sign of puberty is a growth spurt (Archibald, Graber, & Brooks-Gunn, 2006). During puberty, boys become more muscular than girls, pubic hair begins to grow, and voices change (Patterson, 2008). In addition, motor coordination increases in both sexes. Interestingly, those who are well-off enter puberty earlier than do those who are impoverished (Eveleth & Tanner, 1990). These differences in development may occur because those who are impoverished are also more likely to suffer from malnutrition.

Cognitive and Neurological Development

Early research on adolescence done by Piaget suggests adolescents enter what Piaget dubbed the formal operational stage. During this stage of development, adolescents can use abstract thought and represent items using symbols (Piaget, 2000). More recent research has examined the neurological correlates behind cognitive development in adolescence. For example, as children enter adolescence, they become more efficient processors of information. This increased processing speed and efficiency are in part due to a pruning of synapses that leads to more efficient processing (Keating, 2004) as well as increases in myelination that lead to increased processing speeds (Webb, Monk, & Nelson, 2001).

In fact, some neurological development may also be associated with changes in emotional reactivity in adolescents. Sensitivity among neurotransmitters changes as adolescents develop, and this increased sensitivity may be associated with some of the emotional instability during adolescence. Such emotional instability can be expressed as disproportionate reactions to events, for example, reacting overly pessimistically to negative events while being overly elated in reaction to positive events (Archibald et al., 2006). The changes in hormone sensitivity may also be associated with the changes in rates of depression in adolescence. Although rates of depression are similar between boys and girls before adolescence, after entering adolescence girls are more likely to be depressed than are boys (Cicchetti & Toth, 1998). Throughout adolescence, emotional capabilities are greater than in childhood; however, these greater complexities can also lead to struggles in adolescence.

Emotional Development

Although the hormones released during adolescence lead to the physical development associated with puberty, the release of these hormones may also be associated with emotional instability in adolescence. Additionally, emotional relationships become more complex in adolescence (Lane & Schwartz, 1987). Adolescents begin to spend more time with friends than they do with their parents. Especially for those adolescents who enter puberty at an earlier age, a large part of this increased emotional complexity is due to greater amounts of interaction between the sexes, often leading to romantic relationships (Friedlander, Connolly, Pepler, & Craig, 2007). Adolescents begin to explore their burgeoning sexuality as mixed-sex cliques become more common (Patterson, 2008).

Physical Vulnerability

Early Physical Condition and Early Physical Environment

The effects of economic hardship on development affect both physical and mental health outcomes among adolescents and young adults. The negative effects of socioeconomic hardship are especially noticeable when poverty is extreme and persistent or occurs early in a person's life as these types of poverty are more likely to be related to negative developmental outcomes compared with transient poverty (McLeod & Shanahan, 1996). In general, those living in low socioeconomic status neighborhoods have more physical and mental health problems (Chappell & Funk, 2010; Flouri, Tzavidis, & Kallis, 2010) and are less likely to exercise (Cleland et al., 2010) possibly leading to higher levels of obesity than those in higher socioeconomic status neighborhoods. Over the life course, growing up as a child in poverty has long-term life consequences as an adult, such as an increased rate of suicide (Rojas & Stenberg, 2010), poorer health outcomes due to low-quality environmental conditions during key developmental periods in life (Conroy, Sandel, & Zuckerman, 2010), and a lower life expectancy compared with those who grew up in a higher socioeconomic status group (Clarke et al., 2010). Adolescents from low-income families seem to be at higher risk for unsafe sexual behaviors, in part due to higher HIV prevalence rates in urban settings and greater availability of illicit drugs (Tubman, Windle, & Windle, 1996). Accordingly, youth living in poverty are at greater risk for HIV exposure in view of their higher rates of STD infection and teenage pregnancies (Romer et al., 1994).

Research suggests that poor-quality physical environments—often associated with poverty—have detrimental and long-term effects on adolescents. As early as ages 2 and 3, children who grow up with more of these poor conditions such as smoking in the home, maternal anxiety, and exposure to violence are more likely to be diagnosed with asthma than are those children who grow up with better quality

conditions (Berz et al., 2007). Those children diagnosed with asthma feel the stress associated with this chronic illness into childhood and early adolescence as those youth in urban disadvantaged areas had poor quality and little amount of knowledge about asthma as well as how to treat and control their asthma (Mitchell & Murdock, 2005). Being born into a low socioeconomic status environment is also associated with a higher risk of psychiatric symptoms, including autism, attention deficit hyperactivity disorder, and internalizing symptomatology, in adolescents who had a low birth weight. Low birth weight and premature birth are risk factors in and of themselves as both are also associated with negative psychiatric outcomes (Indredavik et al., 2010).

Sleep Problems

Sleep patterns and sleep problems seem to be an important health factor that is related to physical vulnerability in adolescence and young adulthood. Sleep deprivation in adolescence is associated with feelings of sleepiness, irritation, and loss of attention and behavior regulation (Beebe et al., 2008). Feelings of sleepiness are associated with high levels of anxiety and depression and low levels of perceived health (Moore et al., 2009). Sleeping too little or too much is also associated with adolescent health outcomes. For example, in Taiwanese adolescents, sleeping for short periods was related to higher levels of health-risk behaviors such as suicidality, violence, drug use, unprotected sex, and delinquent behaviors, and sleeping for extended periods was also related to higher levels of these behaviors except for suicidality (Yen, King, & Tang, 2010). These studies suggest that not only is sleep a vital process needed for functioning, but sleep is also a protective factor helping to inhibit risky behaviors. However, too little or too much sleep may be a sign of an underlying psychological or physical health problem.

Diet, Lifestyle, and Stress

Another environmental and health factor to consider regarding physical vulnerability among adolescents and young adults is diet and nutrition intake. Those adolescents who grow up in poverty are more likely to be "food insecure" or not to be certain that they will have enough to eat consistently on a daily basis. Adolescents who are food insecure are more likely to perceive eating healthy as inconvenient and believe that healthy food does not taste as good as unhealthy food. Food insecure youth are also more likely to eat fast food and have a higher fat intake than those adolescents who are food secure (Widome, Neumark-Sztainer, Hannan, Haines, & Story, 2009). For adolescents, a diet consisting largely of takeout foods, high-sugar treats, and large amounts of red meat is associated with higher levels of social withdrawal and depression as well as higher levels of delinquency.

In contrast, a diet rich in green, leafy vegetables and fruits was associated with low levels of these negative outcomes (Oddy et al., 2009). Quality of diet is also associated with some mental health outcomes such as bipolar disorder (Jacka et al., 2011) and depression (Jacka et al., 2010) such that those adolescents with poorer diets also have higher rates of mental illnesses. Indeed, this relationship exists even after controlling for socioeconomic characteristics, family education, and family cohesion. It should be also noted that the direction of the relationship between diet and mental health outcomes, as well as any causality, is not yet known. In general, poverty, poor-quality neighborhood conditions, and poor diet—often concurrent—are associated with increased stressors and poor outcomes for adolescents and young adults who experience them.

When examining substance use in adolescents and young adults, those adolescents with an earlier onset of smoking, higher levels of smoking, and faster increases in levels of substance use are more likely to become dependent as young adults compared with those adolescents who began using substances later in life, used substances less frequently, or stopped using substances (Van De Ven, Greenwood, Engels, Olsson, & Patton, 2010). Substance use can also be associated with life stressors. For example, in a study examining risk of substance use in a national sample of 12- to 17-year-olds, adolescents who had experienced physical or sexual abuse were more likely to use and abuse substances. This effect was particularly strong for boys who were more likely to use or abuse alcohol and marijuana compared with girls (Kilpatrick et al., 2000). Similarly, using data from the Rochester Youth Development Study, researchers found that exposure to intimate partner violence in adolescence was associated with higher levels of substance use in young adulthood (Smith, Elwyn, Ireland, & Thornberry, 2010). However, this effect was found only for girls and not for boys. As illustrated by the results of these studies, there seems to be significant gender differences in the effects of diverse stressors on substance use behaviors.

Of use for examining substance use prevention, evidence suggests that engaging in a physical activity, such as a sport, was associated with lower levels of substance use in adolescents. Not only did physical activity help prevent substance use, but membership in a sports team also had protective benefits (Elder, Barnes, Leaver-Dunn, Nagy, & Leeper, 2001). Therefore, such findings suggest that belonging to a group may be just as important to preventing use of alcohol, tobacco, and marijuana as the physical activity that group provides.

Disability and Chronic Illness

Coping with a disability or a chronic illness may be difficult for adolescents and their families. For example, adolescents who feel a great deal of anxiety regarding pain or who are very vigilant of pain are more likely to categorize their pain levels as high intensity after surgery than are those adolescents who feel low levels of anxiety and vigilance regarding pain (Lautenbacher et al., 2010). In addition,

families of children with disabilities also experience heightened stress and must learn to cope with the difficulties associated with their child's disability. However, mother's appraisal of the child's disability and positive affect as a coping resource can help families cope. Those mothers with high levels of positive affect had families that saw the positive side of childhood disability and maintained higher levels of positive emotions (Trute, Benzies, Worthington, Reddon, & Moore, 2010).

Cognitive Vulnerability

Brain and Neurological Development in Adolescence and Early Adulthood

Recent research in developmental neuroscience indicates the role of adolescents' brains on their risk-taking decision-making and suggests that development of the frontal lobes leads to better self-regulation as well as better risk assessment; yet, the combination of relatively higher inclination to seek rewards and still maturing capacities for impulse control may contribute to heightened vulnerability to make more risk-taking decisions during adolescence (Steinberg, 2008, 2010).

Increasing evidence from human imaging and animal studies suggests that on a neurobiological level, risk-taking in adolescence seems to derive, in part, from the distinct neurobiological bases and developmental trajectories associated with two neural systems that underlie impulse control and risky decisions (Casey, Getz, & Galvan, 2008; Galvan, 2010; Steinberg, 2010): a network sensitive to rewards and risks and a network associated with cognitive control. Impulsivity diminishes with age across childhood and adolescence and is associated with protracted development of the prefrontal cortex, but risk-taking increases during adolescence relative to childhood and adulthood and is associated with subcortical systems known to be involved in evaluation of rewards (Casey et al., 2008). Adolescence is a critical developmental period for each system, and dysfunction in either or both networks may contribute to risk-taking behaviors.

The neural network sensitive to reward, especially regions of the striatum, is stimulated by a dramatic and rapid increase in dopaminergic activity occurring during adolescence (Bickel et al., 2007; Steinberg, 2008; Wahlstrom, White, & Luciana, 2010). It is hypothesized that cortical pruning may lead to diminished dopamine autoreceptor function that is critical for inhibitory control of dopamine release (Steinberg, 2010; Wahlstrom et al., 2010). Recent studies using functional magnetic resonance imaging (fMRI) provide converging evidence for the hypothesis that adolescents are hyperresponsive to rewards. In particular, studies of reward processing indicate that regions of the striatum that receive strong projections from midbrain dopamine neurons are more active in adolescents relative to adults in response to and anticipation of rewards (Ernst et al., 2005; Galvan, 2010; Galvan

et al., 2006). These studies suggest that the pursuit of high-risk yet rewarding options among adolescents may be driven by neural processes that are sensitive to rewards rather than their inability to estimate consequence probability or an attitude of excessive invulnerability (Boyer, 2006).

A neural network that includes regions of the prefrontal cortex—associated with cognitive control—undergoes maturation throughout adolescence and into early adulthood (Paus, 2005). Increased myelination along with experience-dependent synaptogenesis and pruning is believed to lead to the functional maturity of neural circuits underlying cognitive control. The neural network involving the prefrontal cortex is hypothesized to interact with the reward system to contribute to impulse control difficulties in adolescence (Bickel et al., 2007; Casey et al., 1997, 2008). Studies involving functional neuroimaging of tasks that require an ability to overcome interference from multiple stimuli or prepotent response biases demonstrate a developmental pattern in which task-related activations are initially diffuse among prefrontal regions in childhood or early adolescence (Casey et al., 1997; Casey, Giedd, & Thomas, 2000), but the diffuse patterns of activity become more focal with maturation. As a result, both cognitive control performance and neural patterns of activity come to resemble adult performance and neural patterns with increasing age.

Recent work by Bickel et al. (2007) involves delayed discounting of rewards to examine the relationships between cognitive control and reward systems. Delay discounting refers to the decline in the valuation of a reward as a function of the delay in receiving it and is measured by the preference for either a smaller, more immediate amount or a delayed but larger amount (Madden & Bickel, 2009). Neuroimaging studies indicate that delay discounting recruits two distinct neural systems: one limbic and one prefrontal. Specifically, the limbic system (midbrain dopamine system including paralimbic cortex, the network associated with reward sensitivity) is preferentially activated for decisions involving immediately available rewards, whereas the lateral prefrontal cortex and posterior parietal cortex (the network associated with cognitive control) are engaged for decisions involving delayed rewards (McClure, Ericson, Laibson, Loewenstein, & Cohen, 2007; McClure, Laibson, Loewenstein, & Cohen, 2004).

Delay discounting discriminates between individuals with and without addictive disorders, drug use problems, poor health behaviors, and risky sexual behaviors (Bickel et al., 2010; Madden & Bickel, 2009; Wilson & Daly, 2004). In a study that compared discounting of heroin-dependent individuals who either did or did not share injecting equipment, heroin addicts who shared equipment discounted money more than heroin addicts who did not share injection equipment, demonstrating that monetary discounting can distinguish risk-taking behaviors (Odum, Madden, Badger, & Bickel, 2000). In a sample of adolescents and young adults, monetary discounting was significantly associated with a variety of risky sexual behaviors (Chesson et al., 2006). Taken together, the studies of brain and neurological development indicate that the interaction of reward and cognitive control systems predicts risk-taking behaviors in adolescence and young adulthood.

There is evidence emphasizing the role of working memory (WM) in risk-taking decisions and behaviors as indicated by the negative correlation between delay discounting and WM. WM is the ability to maintain active representations of goal-relevant information despite interference from competing or irrelevant information (Engle, 2002). For example, in a young adult sample WM-related neural activity in the left anterior prefrontal cortex covaried with delay discounting ($r = -0.40$) and partially mediated the relation between intelligence and delay discounting (Shamosh et al., 2008). This result suggests that delay discounting is associated with intelligence in part because of processes instantiated in the anterior prefrontal cortex, a region known to support the integration of diverse information. This finding further explains the tendency of more intelligent individuals to resist smaller, immediate rewards. Similarly, in a study based on a sample of young adults that included participants with and without a history of alcohol dependence, lower WM capacity was related to higher delay discounting, which seems to reflect a general vulnerability to externalizing, disinhibitory disorders (e.g., lifetime problem with drug use, childhood conduct problems, and adult antisocial behavior problems) (Bobova, Finn, Rickert, & Lucas, 2009). Furthermore, a recent study by Bickel and colleagues showed that neurocognitive training of WM resulted in a decrease in discounting of delayed rewards among stimulant-dependent individuals (Bickel, Yi, Landes, Hill, & Baxter, 2011). Such a finding corroborates the abovementioned studies supporting a strong relationship between WM and delay discounting and previous studies using fMRI indicating an overlap of brain regions that are activated during WM and delay discounting tasks (McClure et al., 2004, 2007; Owen, McMillan, Laird, & Bullmore, 2005).

The Role of Effortful Control in the Development of Health-Risk Behaviors

A key step to building knowledge about the etiology of risk-taking and health-risk behaviors in adolescence and young adulthood is to examine the role of personality systems including specific behaviors rooted in temperament that arise early in childhood. Temperament includes behaviors that vary widely across individuals, are readily observed early in life, are somewhat stable over time and across settings, and include a biological foundation (Buss & Plomin, 1984; Goldsmith, Buss, & Lemery, 1997; Rothbart & Bates, 1998; Sanson, Hemphill, & Smart, 2004; Strelau, Zawadzki, & Piotrowska, 2001). Individual differences in temperament are behavioral manifestations of underlying brain mechanisms involved in energy, response to new information (specifically, tendency to approach and withdraw), and regulation of arousal (Nęcka, 2003). To understand the contribution of temperament/personality factors to risk behaviors among adolescents and young adults, we focus on effortful control, one of the three facets of temperament according to Rothbart's theory (Rothbart & Bates, 2006).

Volitional control of attention, or effortful control, is critical to the regulation of reactive emotion and behavioral/motor states (Posner & Rothbart, 2000, 2006). Higher levels of effortful control of attention are associated with lower levels of frustration/anger and surgency/extraversion, with attention serving as the key cognitive/behavioral regulatory mechanism (Calkins & Fox, 2002; Eisenberg et al., 2005; Kochanska & Knaack, 2003; Olson, Sameroff, Kerr, Lopez, & Wellman, 2005; Rothbart, Ellis, Rueda, & Posner, 2003; Rydell, Berlin, & Bohlin, 2003). Recent theories suggest that, in childhood and adolescence, those who are more reactively aroused (such as those who are more readily activated by stimulation) and who have a more difficult time regulating that arousal are more likely to show risk-taking problems (Strelau et al., 2001).

With respect to attention serving cognitive self-regulation, a recent longitudinal study based on the participants from the Study of Early Child Care and Youth Development assessed throughout middle childhood and early adolescence shows that levels of inattention moderated the link between negative emotionality and externalizing problems (Kim & Deater-Deckard, 2011). Specifically, for those with poor attention control, increases in anger predicted larger increases in externalizing problems over time. For those with good attention control, the lagged predictive effects of anger on externalizing problems notably decreased over time resulting in these effects becoming nonsignificant by early adolescence. Thus, the developmental link between changes in anger and changes in externalizing problems was more obvious and persistent over time for individuals with poor attention regulation—a longitudinal effect that is consistent with prior shorter term longitudinal and cross-sectional studies (Eisenberg et al., 2004, 2009; Lawson & Ruff, 2004; Oldehinkel, Hartman, Ferdinand, Verhulst, & Ormel, 2007).

Adolescents and young adults who struggle with self-regulation are more likely to engage in risky behaviors because they are less able to restrain responses to temptation (Brynes, 1998; Steinberg, 2004). Indeed, empirical evidence demonstrates that poor self-regulation of impulsivity is related to risk-taking behaviors (Auerbach, Abela, & Ho, 2007; Breakwell, 1996; Caspi et al., 1997; Hoyle, Fejfar, & Miller, 2000; Krueger, Caspi, & Moffitt, 2000; Moilanen, Crockett, Raffaelli, & Jones, 2010). Furthermore, human and animal survey and neuroscience literature point to the role of frontal cortex activation in the regulation of activity in more primitive limbic regions involved with approach and avoidance impulses and affects (Davidson, Jackson, & Kalin, 2000; Gray, 2004; Heilman, 1997; Posner & Rothbart, 2005). Overall, extant literature suggests that effective self-regulation of negative emotions and impulsive behaviors may be fundamental to risk-taking behaviors among adolescents and young adults (Blair & Diamond, 2008; Kim & Deater-Deckard, 2011; Quay, 1993; Rothbart & Bates, 2006; Steinberg, 2008).

Negative Attributional Style as a Vulnerability to Emotional Problems

In recent years, several cognitive vulnerability models have been put forth to account for the onset and maintenance of depression. These models posit that depression is associated with, and maintained by, preexisting biased information processes focused on themes of personal failure and loss. Early on, Beck (1967) viewed the measurement of maladaptive self-schemas as one way to capture the cognitive vulnerabilities associated with depression. Subsequently, in their reformulation of the learned helplessness model, Abramson, Seligman, and Teasdale (1978) proposed that maladaptive attributional styles constituted a cognitive vulnerability and further suggested that these styles were pivotal in the onset and maintenance of depression. Specifically, a cognitive style that attributed negative events to internal, stable, and global causes and positive events to external, unstable, and specific causes was associated with depression. In contrast, the perception of positive events as being due to internal, stable, and global causes and negative events as being due to external, unstable, and specific causes might buffer or protect individuals from the development of depression. Abramson, Metalsky, and Alloy (1989) reformulated the helplessness model into what has come to be called the hopelessness model, proposing that negative life events may foster negative inferences about the self involving a sense of futility and hopelessness, which in turn, serves as a proximal contributory cause of depressive symptoms.

Clearly, attributional style is a cognitive vulnerability factor that has gained considerable theoretical and research attention in the literature. A number of prospective studies have been conducted that link negative attributional style with onset of depression in youth, though the evidence is more consistent with adolescents (and adults) than it is with children (Hankin, Lakdawalla, Carter, Abela, & Adams, 2007). From a developmental perspective, emerging evidence concerning younger children indicates that attributional style mediates, rather than moderates, the associations between life stress and depression (Cole et al., 2008; Gibb & Alloy, 2006; Hankin, 2008; Hankin & Abramson, 2001; Kim-Spoon, Ollendick, & Seligman, 2012). That is, negative attributional style is thought to result from an accumulation of negative life events and to play a pivotal role in subsequent development of depression. For older children and adolescents, however, negative attributional style is thought to be already developed and to be more stable. This more stable depressive attributional style may serve as a risk factor for depression among youth (Cole et al., 2008; Gibb & Alloy, 2006).

In a recent longitudinal study, Kim-Spoon et al. (2012) examined the interactive effects of depressive attributional style and multiple domains of perceived competence on depressive symptoms among adolescents who were assessed three times in the fall of the eighth grade (Time 1), the spring of the eighth grade (Time 2), and the spring of the tenth grade (Time 3). Their results based on structural equation modeling analyses revealed significant interaction effects between attributional style and perceived competence such that for girls with a higher depressive

attributional style, lower perceived competence in physical appearance was predictive of depressive symptoms over a 2.5-year period (even after controlling for previous levels of depressive symptoms measured 2 years earlier). Regardless of gender, among adolescents with a higher depressive attributional style, lower athletic competence was predictive of higher depressive symptoms 6 months later, which in turn were related to higher depressive symptoms 2 years later. In addition, gender differences in main effects suggested that lower levels of perceived social acceptance were associated with higher subsequent levels of depressive symptoms but only for boys. These findings illustrate how depressive attributional styles interact with certain dimensions of perceived competence (particularly athletic abilities and physical appearance) to contribute to vulnerability to depression among adolescent boys and girls.

Emotional Vulnerability

Emotion Reactivity and Emotion Regulation

The understanding, expression, and regulation of emotions change across the lifespan, and these changes are particularly evident in adolescence. In adolescence, many physical changes take place that have implications for changes in emotion development. Emotions during adolescence, in part due to the influx of hormones, are characterized by overall increased intensity—in both positive and negative emotional states. As children enter puberty, there is an increase in multiple hormones including hormones such as thyroxine, adrenal androgen, testosterone, and estrogen. Not surprisingly, many developmental problems of adolescence tend to involve emotional extremes and poor emotional control. As a result, adolescents are at an increased risk of depression, suicide, and delinquency. In particular, emotions can influence risk-taking behaviors and decision-making processes in two broad ways (Boyer, 2006).

First, how people react to emotion-provoking experiences may influence decision-making in potentially risky situations. The key idea of the somatic marker hypothesis is that decision-making is a process that is influenced by marker signals that arise in bioregulatory processes, including those that express themselves in emotions and feelings (Bechara & Damasio, 2005; Bechara, Damasio, & Damasio, 2000; Bechara, Damasio, Damasio, & Lee, 1999). Thus, emotional responses to positive and negative consequences guide decision-making in risky and uncertain situations. Furthermore, research has demonstrated an interesting interaction between neurocognitive performance and emotional distress among substance-dependent, HIV-positive individuals suggesting that greater emotional distress (such as depression and anxiety) was predictive of greater sexual risk behaviors only among those who showed better performance in the Iowa Gambling Task (Wardle, Gonzalez, Bechara, & Martin-Thormeyer, 2011). Such a result suggests

that a neural circuitry that is critical for affective decision-making needs to be intact for emotional distress to increase risk-taking behaviors.

Second, emotion regulation, which relates to the control of affect, drive, and motivation, may influence the degree of risk-taking propensity. Emotionally dysregulated adolescents and young adults (particularly those who are impulsive or anger prone) are inclined toward risk-taking behaviors including substance use and risky sexual behaviors (Cooper, Wood, Orcutt, & Albino, 2003; Eisenberg et al., 2005; Magar, Phillips, & Hoise, 2008). Recent research suggests the potential importance of cognitive strategies for emotion regulation in curbing risk-seeking behaviors. For example, effective use of emotion regulation strategies (e.g., cognitive reappraisal) attenuate risky decision-making and modulate neural responses associated with reward processing or risk aversion (Heilman, Crisan, Houser, Miclea, & Miu, 2010; Martin & Delgado, 2011).

Research evidence has accumulated suggesting that emotion regulation and subsequent socio-emotional development play important roles in both psychological health and physiological health. For instance, adolescents who reported more intense and labile emotions and less effective regulation of these emotions also reported more depressive symptoms (Kovacs, Joormann, & Gotlib, 2008). Furthermore, good emotion regulation capabilities can act as a protective factor against trauma and other risk factors, whereas poor emotion regulation may act as a risk factor leading to more negative emotions, poor coping, and poor social interactions. In particular, emotion regulation may act as adaptive factors in the face of adversity and may promote more positive coping strategies and thus healthy adjustment throughout development. For example, research has shown that children's optimal emotion regulation is negatively affected by maltreatment and inter-adult violence in early and middle childhood (Kim & Cicchetti, 2010; Maughan & Cicchetti, 2002) and adolescents' poor emotion regulation is related to higher levels of anxious/depressed symptoms and social problems (Lawson & Ruff, 2004; Silk, Steinberg, & Morris, 2003). Overall, past research points out that emotion regulation can both directly and indirectly affect emotional adjustment among adolescents, through its influences on the ways in which youth process and control their reactions to their environments.

Dual Motivation-Emotion Pathways to Risk-Taking Behaviors

Traditional theories of risk-taking have heavily focused on the regulation of arousal and positive emotions (Zuckerman, 1983). According to this perspective, risk-taking is viewed as an appetitive or approach behavior in which individuals choose risky options as a way to achieve optimum arousal, enhance positive emotional experience, or both (Cooper, Flanagan, Talley, & Micheas, 2006). Reward or sensation seeking is defined as stable individual differences in preferences for varied, novel, complex, and intense sensations and experiences, and such individual differences are thought to be rooted in the behavioral activation system (BAS) (Zuckerman, 1994).

Indeed, young adults with higher levels of reward/sensation seeking are more likely to be engaged in a variety of risk behaviors including illicit drug use, heavy drinking, and risky sexual behaviors (Katz, Fromme, & D'Amico, 2000).

Cooper et al. (2006) proposed the idea that distinct positive and negative motivation-emotion pathways underlie risk-taking behaviors. This view is broadly consistent with major motivational theories of behavior including Gray's (1970) two neurologically distinct motivation systems that underlie behavior. Specifically, the behavioral inhibition system (BIS) regulates behaviors that involve avoidance or escape from negative or painful experiences, whereas BAS regulates behaviors that involve the pursuit of positive or pleasurable ones. Empirical evidence supports the idea of dual pathways to risk-taking, showing that people attribute both positive-mood-enhancing and negative-mood-altering propensities to a wide range of risky behaviors.

Therefore, risk-taking behaviors may be the product of a person's effort to downregulate emotional distress rather than the experience of emotional distress per se. In general, researchers agree on the crucial distinction between two types of emotion regulation strategies: strategies to regulate one's negative emotions through active approach or engagement (e.g., analyzing the problem) versus strategies that involve avoidance or withdrawal from the stressful situations or denial of, minimization of, or escape from the emotions themselves (Thayer, Newman, & McClain, 1994). Research shows that avoidant strategies are not only less effective ways to regulate negative emotions (Cooper et al., 2006) but are also consistently associated with diverse risk-taking behaviors, including alcohol use and drinking problems (Cooper, Frone, Russell, & Mudar, 1995), externalizing behaviors (Loesel & Bliesener, 1994), and risky sexual behaviors (Cooper, Agocha, & Sheldon, 2000; Cooper, Shapiro, & Powers, 1998; Folkman, Chesney, Pollack, & Phillips, 1992; Schachner & Shaver, 2004).

As mentioned earlier, heightened levels of reward and sensation seeking during adolescence may reflect the pattern of brain development during this developmental period. Specifically, the frontal regions of the cerebral cortex increase their connections and integrated activity with the temporal lobes during middle childhood indicating greater neural efficiency. The brain undergoes a growth spurt in the frontal lobes of the cerebral cortex in late adolescence, around age 17, and this growth continues into early adulthood. This area of the brain is in control of logic, planning, and emotion regulation. Therefore, adolescence is characterized by slowly developing better control of emotions and their expression. Adolescents, with the frontal lobes not having achieved full development, are much more likely to engage in sensation-seeking and risky behaviors.

Sex and Ethnic Differences Vulnerabilities

Research examining sex differences indicates that boys hold riskier health behavior beliefs and more frequently engage in health-risk behaviors than do girls (Courtenay, McCreary, & Merighi, 2002). These beliefs and behaviors are held in

regard to things such as diet, substance use, medical care, and amount of sleep. Compared with girls, boys are more likely to eat a low-nutrient, high-fat diet; have higher rates of substance use and abuse; and have higher rates of noncompliance with medical recommendations (Courtenay et al., 2002). Female adolescents have also been found to get more sleep than male adolescents (Moore et al., 2011). The poor health behaviors demonstrated by boys may be associated with their lower longevity as well as higher rates of some illnesses compared with girls.

In adolescence and young adulthood, boys seem to engage in risk-taking behaviors more frequently than do girls (Rosenblitt, Soler, Johnson, & Quadagno, 2001). Although, there is a great paucity of information available regarding sex differences in the predictors of risk behaviors, research indicates some evidence of sex differences in the neural process of risk-taking. An fMRI study of young adults showed that, when taking the same level of risk, young women showed stronger activation in the right insula and bilateral orbitofrontal cortex (OFC) than did young men (Lee, Chan, Leung, Fox, & Gao, 2009). In addition, the strength of neural activity in the insula was correlated with the rate of risky behaviors for girls but not for boys. Similarly, the percent signal changes in the right OFC were correlated negatively with the rate of selecting risky choices for girls only. These findings suggest a sex-related influence modulating brain activity during risk-taking tasks.

A recent meta-analytic study on sex differences in impulsivity demonstrated that the observation of male subjects showing greater rates of risk-taking behaviors is likely explained by sex differences in punishment sensitivity rather than sex differences in reward sensitivity and effortful control (Cross, Copping, & Campbell, 2011). Specifically, female subjects were consistently more punishment sensitive, but male subjects did not show greater reward sensitivity. Though male subjects showed significantly higher sensation seeking, sex differences were not found on delay discounting or executive function tasks. Therefore, these findings suggest a stronger sex difference in motivational rather than effortful or executive forms of behavior control.

Sex differences have also been observed in levels of emotion regulation, indicating that girls show better emotional regulation skills than boys do (Silk et al., 2003). Furthermore, some studies have found effects of emotion regulation to be effective in preventing unhealthy adjustment or risk behaviors in boys but not girls. In intervention studies targeting inner-city African American youth, social-emotional programs that emphasized communication, social networking, stress management, empathy, and role models were effective in preventing boys from engaging in multiple risk behaviors, including violence. However, this effect was not found for girls (Flay et al., 2004). Thus, preliminary findings from extant literature seem to suggest that boys show lower levels of emotion regulation than girls do; however, emotional regulation appears to play a more important role as a protective factor for adjustment problems among boys.

Some studies suggest sex differences in emotional regulation among late adolescents and young adults (Domes et al., 2010; McRae, Ochsner, Mauss, Gabrieli, & Gross, 2008). Specifically, in these studies, fMRI signals from the prefrontal cortex regions of the brain were used to identify sex differences in control-related

regions that were more active during cognitive regulation when compared with responses to negative emotion pictures. The findings indicated that male subjects showed greater downregulation (i.e., cognitive appraisal involving changing the meaning of the event to decrease its emotional impact) and less activity in prefrontal regions than female subjects responding to negative images. Therefore, although male and female subjects do not report differences in the frequency with which they use cognitive emotion regulation in everyday life, when they are instructed to regulate, male subjects seem to be able to regulate their negative emotion with greater efficiency or less effort than female subjects. Specifically, these studies suggest that sex difference in emotional responding might be attributed to enhanced emotional reactivity, reduced capabilities to cognitively regulate in terms of reappraisal, or both.

There may also be some differences between the sexes such that interventions and protective factors operate differently. For example, interventions to decrease substance use may be more effective for female than for male adolescents (Godley, Hedges, & Hunter, 2011). In addition, some Head Start interventions—specifically those aimed at children in poverty—show lasting benefits into adolescence for girls. For girls, this intervention improved grades and lowered rates of suspension. On the other hand, this intervention was unrelated to outcomes for boys (Joo, 2010). However, while boys may not be as receptive to interventions as are girls, the family may provide more protection from negative environmental influences, such as neighborhood violence, for male than for female children and adolescents (McKelvey et al., 2011).

In terms of ethnic differences in physical vulnerability factors among adolescents and young adults, nonminority adolescents have been found to get more sleep as well as more regular sleep than do minority adolescents, which may be linked to important health outcomes (Moore et al., 2011). There seem to be some ethnic differences in the predictive factors that are associated with health outcomes such as obesity. For Latinas and Asian boys as well as Latino and White boys, low parental education was found to be a risk factor for being overweight. However, for White girls and boys, watching more than 2 h of television or playing more than 2 h of video games was associated with being overweight (Wilkosz, Chen, Kennedy, & Rankin, 2010). Additionally, there may be some ethnic differences associated with physical environment. For example, not only are those of an ethnic minority more likely to live in poverty, they are also more likely to be exposed to severe violence than are those of ethnic majority (Selner-O'Hagan, Kindlon, Buka, Raudenbush, & Earls, 1998).

Although evidence suggests that ethnic minority adolescents (particularly African Americans) are more likely than are White adolescents to engage in risk-taking behaviors such as risky sexual behaviors (Blum et al., 2000; Santelli, Lowry, Brener, & Robin, 2000), little research has explored or found meaningful ethnic differences in the predictive power of neurological, cognitive, and emotional vulnerability factors for adolescent risk behaviors (Wallace, Brown, Bachman, & Laveist, 2003). Nevertheless, available preliminary findings indicate significant ethnic differences in some of the cognitive and emotional vulnerability factors.

First, African American and Latino youth were both rated as demonstrating more severe symptoms on the observational ratings of depression compared with Caucasian youth. In terms of cognitive variables associated with depression, African Americans reported fewer negative cognitive biases (including negative attributional style) compared with Caucasians; but cognitive biases were significantly correlated with depression severity across all ethnic groups (Stein et al., 2010). Second, ethnic differences in emotion regulation strategies indicated that ethnic minority young adults used suppression to regulate emotion more often than European American young adults (Gross & John, 2003).

Conclusion

The studies discussed in this chapter indicate that diverse physical, cognitive, and emotional vulnerability factors play crucial roles in physical and psychological health by influencing risk behaviors during adolescence and young adulthood. There are sex differences in risk beliefs and behaviors such as substance use and diet, although it is important to emphasize that there also is a great deal of similarity between boys and girls of vulnerability factors that influence their risk behaviors. The physical environment and condition in which children and adolescents grow up have significant long-term effects on physical and psychological health outcomes. In particular, poor-quality physical environments that are often associated with poverty seem to be related to higher vulnerabilities to psychiatric symptoms and health-risk behaviors such as use of illicit drugs and risky sexual behaviors. Sleep and diet also seem to make significant contributions to physical and mental health outcomes among adolescents and young adults. Cognitive vulnerability factors influence adolescents' and young adults' healthy development, including neural network systems that are related to heightened reward sensitivity and reduced cognitive control, poor WM, low effortful control, and negative attributional styles. Recent research on developmental neuroscience further indicates that delay discounting is strongly related to addictive behaviors, drug abuse, poor health behaviors, and risky sexual behaviors. Finally, emotional vulnerability factors are strongly tied to health and well-being during adolescence and young adulthood. Poor emotion regulation and avoidant coping strategies seem to influence risk-taking propensity and behaviors, resulting in increased risks for developing behavioral and emotional problems and disorders that range from substance use problems to mood disorders such as depression.

Although a multitude of vulnerability factors exists for physical and psychological health among adolescents and young adults, research suggests many protective factors that can be targeted as promising avenues for interventions. These protective factors include individual characteristics such as high scholastic achievement, strong religious identity, and high self-esteem (Blum & Rineheart, 2000; Resnick et al., 1997). As well, for those adolescents who must learn to cope with chronic pain or disability, a potential point of intervention may be learning more effective

coping strategies. Adaptive coping strategies are associated with greater muscle relaxation, whereas maladaptive coping strategies are associated with greater muscle tenseness (van der Hulst, Vollenbroek-Hutten, Schreurs, Rietman, & Hermens, 2010). Finally, research evidence emphasizes the importance of parent-family connectedness and perceived school connectedness as protective factors against adolescents' health-risk behaviors (Resnick et al., 1997). Though adolescents are increasingly becoming independent from their parents in many ways, continuing warm, supportive relationships with parents as well as feeling connected at school are both protective.

The goal for future research includes investigating the specific mechanisms in the interplay between individual's biological and psychological characteristics and environmental influences, coupled with learning and shifts in the context of adolescents and young adults as they grow older. Of particular importance to the advancement of prevention science is to explain why some individuals who appear to be at low risk for health and adjustment problems end up with poor health and maladjustment, whereas other individuals who seem to be at high risk turn out highly resilient. Such work will be invaluable to design prevention and intervention strategies that will reduce health-risk behaviors in adolescence and young adulthood.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory based subtype of depression. *Psychological Review, 96*, 358–372.

Abramson, L. Y., Seligman, M. E. B., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49–74.

Archibald, A., Graber, J., & Brooks-Gunn, J. (2006). Pubertal processes and physiological growth in adolescence. In G. Adams & M. Berzonsky (Eds.), *Blackwell handbook of adolescence*. Malden, MA: Blackwell.

Auerbach, R. P., Abela, J. R. Z., & Ho, M. H. R. (2007). Responding to symptoms of depression and anxiety: Emotion regulation, neuroticism, and engagement in risky behaviors. *Behaviour Research and Therapy, 45*, 2182–2191.

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

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision-making, and the orbitofrontal cortex. *Cerebral Cortex, 10*, 295–307.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience, 19*, 5473–5481.

Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beebe, D. W., Fallone, G., Godiwala, N., Flanigan, M., Martin, D., Schaffner, L., et al. (2008). Feasibility and behavioral effects of an at-home multi-night sleep restriction protocol for adolescents. *Journal of Child Psychology and Psychiatry, 49*, 915–923.

Berz, J. B., Carter, A. S., Wagmiller, R. L., Horwitz, S. M., Murdock, K. K., & Briggs-Gowan, M. (2007). Prevalence and correlate of early onset asthma and wheezing in a healthy birth cohort of 2- to 3-year olds. *Journal of Pediatric Psychology, 32*, 154–166.

Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F., & Baxter, C. (2011). Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biological Psychiatry, 69*, 260–265.

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*, S85–S91.

Bickel, W. K., Jones, B. A., Landes, R. D., Christensen, D. R., Jackson, L., & Mancino, M. (2010). Hypothetical intertemporal choice and real economic behavior: Delay discounting predicts voucher redemptions during contingency-management procedures. *Experimental and Clinical Psychopharmacology, 18*, 546–552.

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*, 899–911.

Blum, R. W., & Nelson-Mmari, K. (2004). The health of young people in a global context. *Journal of Adolescent Health, 35*, 402–418.

Blum, R. W., & Rineheart, P. (2000). *Reducing the risk: Connections that make a difference in the lives of youth*. Minneapolis, MN: Division of General Pediatrics and Adolescent Health, University of Minnesota.

Blum, R. W., Beuhring, T., Shew, M. L., Beawringer, L. H., Sieving, R. E., & Resnick, M. D. (2000). The effects of race/ethnicity, income, and family structure on adolescent risk behaviors. *American Journal of Public Health, 90*, 1879–1884.

Bobova, L., Finn, P. R., Rickert, M. E., & Lucas, J. (2009). Disinhibitory psychopathology and delay discounting in alcohol dependence: Personality and cognitive correlates. *Experimental and Clinical Psychopharmacology, 17*, 51–61.

Boyer, T. W. (2006). The development of risk-taking: A multi-perspective review. *Developmental Review, 26*, 297–345.

Breakwell, G. M. (1996). Risk estimation and sexual behavior: A longitudinal study of 16- to 21-year-olds. *Journal of Health Psychology, 1*, 79–91.

Brynes, J. P. (1998). *The nature and development of decision making: A self-regulation model*. Mahwah, NJ: Lawrence Erlbaum Associates.

Buss, A. H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.

Calkins, S. D., & Fox, N. A. (2002). Self-regulatory processes in early personality development: A multilevel approach to the study of childhood social withdrawal and aggression. *Development and Psychopathology, 14*, 477–498.

Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review, 28*, 62–77.

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology, 54*, 241–257.

Casey, B. J., Castellanos, F. X., Giedd, J. N., Marsh, W. L., Hamburger, S. D., Schubert, A. B., et al. (1997). Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 374–383.

Caspi, A., Begg, D., Dickson, N., Harrington, H., Langley, J., Moffitt, T. E., et al. (1997). Personality differences predict health-risk behaviors in young adulthood: Evidence from a longitudinal study. *Journal of Personality and Social Psychology, 73*, 1052–1063.

Centers for Disease Control and Prevention (CDC). (2010). *HIV surveillance report: Diagnoses of HIV infection and AIDS in the United States and dependent areas, 2009* (Vol. 21). Atlanta, GA: Division of HIV/AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, CDC, U.S. Department of Health and Human Services.

Chappell, N. L., & Funk, L. M. (2010). Social capital: Does it add to the health inequalities debate? *Social Indicators Research, 99*, 357–373.

Chesson, H. W., Leichliter, J. S., Zimet, G. D., Rosenthal, S. L., Bernstein, D. I., & Fife, K. H. (2006). Discount rates and risky sexual behaviors among teenagers and young adults. *Journal of Risk Uncertainty, 32*, 217–230.

Cicchetti, D., & Toth, S. L. (1998). The development of depression in children and adolescents. *American Psychologist, 53*, 221–241.

Clarke, C. A., Miller, T., Chang, E. T., Yin, D., Cockburn, M., & Gomez, S. L. (2010). Racial and social class gradients in life expectancy in contemporary California. *Social Science and Medicine, 70*, 1373–1380.

Cleland, V., Ball, K., Hume, C., Timperio, A., King, A. C., & Crawford, D. (2010). Individual, social, and environmental correlates of physical activity among women living in socioeconomically disadvantaged neighborhoods. *Social Science and Medicine, 70*, 2011–2018.

Cole, D. A., Ciesla, J. A., Dallaire, D. H., Jacquez, F. M., Pineda, A. Q., LaGrange, B., et al. (2008). Emergence of attributional style and its relation to depressive symptoms. *Journal of Abnormal Psychology, 117*, 16–31.

Conroy, K., Sandel, M., & Zuckerman, B. (2010). Poverty grown up: How childhood socioeconomic status impacts adult health. *Journal of Developmental and Behavioral Pediatrics, 31*, 154–160.

Cooper, M. L., Agocha, V. B., & Sheldon, M. S. (2000). A motivational perspective on risky behaviors: The role of personality and affect regulatory processes. *Journal of Personality, 68*, 1059–1088.

Cooper, M. L., Flanagan, M., Talley, A., & Micheas, L. (2006). Individual differences in emotion experience and regulation and their relationships to risk-taking during adolescence. In D. K. Snyder, J. A. Simpson, & J. N. Hughes (Eds.), *Emotion regulation in couples and families: Pathways to dysfunction and health*. Washington, DC: American Psychological Association.

Cooper, M. L., Frone, M. R., Russell, M., & Mudar, P. (1995). Drinking to regulate positive and negative emotions: A motivational model of alcohol use. *Journal of Personality and Social Psychology, 69*, 990–1005.

Cooper, M. L., Shapiro, C. M., & Powers, A. M. (1998). Motivations for sex and sexual behavior among adolescents and young adults: A functional perspective. *Journal of Personality and Social Psychology, 75*, 1528–1558.

Cooper, M. L., Wood, P. K., Orcutt, H. K., & Albino, A. (2003). Personality and the predisposition to engage in risky or problem behaviors during adolescence. *Journal of Personality and Social Psychology, 84*, 390–410.

Courtenay, W. H., McCreary, D. R., & Merighi, J. R. (2002). Gender and ethnic differences in health beliefs and behaviors. *Journal of Health Psychology, 7*, 219–231.

Cross, C. P., Copping, L. T., & Campbell, A. (2011). Sex differences in impulsivity: A meta-analysis. *Psychological Bulletin, 137*, 97–130.

Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. *Annals of the New York Academy of Sciences, 1021*, 1–22.

Davidson, R. J., Jackson, D. C., & Kalin, N. H. (2000). Emotion, plasticity, context, and regulation: Perspectives from affective neuroscience. *Psychological Bulletin, 126*, 890–909.

Domes, G., Schulze, L., Böttger, M., Grossmann, A., Hauenstein, W., Petra, H., et al. (2010). The neural correlates of sex differences in emotional reactivity and emotion regulation. *Human Brain Mapping, 31*, 758–769.

Eisenberg, N., Spinrad, T. L., Fabes, R. A., Reiser, M., Cumberland, A., Shepard, S. A., et al. (2004). The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development, 75*, 25–46.

Eisenberg, N., Sadovsky, A., Spinrad, T. L., Fabes, R. A., Losoya, S. H., Valiente, C., et al. (2005). The relations of problem behavior status to children's negative emotionality, effortful control, and impulsivity: Concurrent relations and prediction of change. *Developmental Psychology, 41*, 193–211.

Eisenberg, N., Valiente, C., Spinrad, T. L., Cumberland, A., Liew, J., Reiser, M., et al. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to

their externalizing, internalizing, and co-occurring behavior problems. *Developmental Psychology, 45*, 988–1008.
Elder, C., Barnes, J., Leaver-Dunn, D., Nagy, S., & Leeper, J. (2001). The influence of group membership on multiple substance use. *Journal of Alcohol and Drug Education, 47*, 19–30.
Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19–23.
Ernst, M., Nelson, E. E., Jazbec, S., McClure, E. B., Monk, C. S., Leibenluft, E., et al. (2005). Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage, 25*, 1279–1291.
Eveleth, P., & Tanner, J. (1990). *World-wide variations in human growth* (2nd ed.). New York: Cambridge University Press.
Flay, B. R., Graumlich, S., Segawa, E., Burns, J. L., Holliday, M. Y., & The Aban Aya Investigators. (2004). Effects of two prevention programs on high-risk behaviors among African American youth: A randomized trial. *Archives of Pediatric and Adolescent Medicine, 158*, 377–384.
Flouri, E., Tzavidis, N., & Kallis, C. (2010). Adverse life events, area socioeconomic disadvantage, and psychopathology and resilience in young children: The importance of risk factors' accumulation and protective factors' specificity. *European Child & Adolescent Psychiatry, 19*, 535–546.
Folkman, S., Chesney, M. A., Pollack, L., & Phillips, C. (1992). Stress, coping, and high-risk sexual behavior. *Health Psychology, 11*, 218–222.
Friedlander, L. J., Connolly, J. A., Pepler, D. J., & Craig, W. M. (2007). Biological, familial, and peer influences on dating in early adolescence. *Archives of Sexual Behaviors, 36*, 821–830.
Galvan, A. (2010). Adolescent development of the reward system. *Frontiers in Human Neuroscience, 4*, 6.
Galvan, A., Hare, T., Parra, C., Penn, J., Voss, H., Glover, G., 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.
Gibb, B. E., & Alloy, L. B. (2006). A prospective test of the hopelessness theory of depression in children. *Journal of Clinical Child and Adolescent Psychology, 35*, 264–274.
Godley, S. H., Hedges, K., & Hunter, B. (2011). Gender and racial differences in treatment process and outcome among participants in the adolescent community reinforcement approach. *Psychology of Addictive Behaviors, 25*, 143–154.
Goldsmith, H. H., Buss, K. A., & Lemery, K. S. (1997). Toddler and childhood temperament: Expanded content, stronger genetic evidence, new evidence for the importance of environment. *Developmental Psychology, 33*, 891–905.
Gray, J. A. (1970). The psychophysiological basis of introversion–extraversion. *Behaviour Research and Therapy, 8*, 249–266.
Gray, J. R. (2004). Integration of emotion and cognitive control. *Current Directions in Psychological Science, 13*, 46–48.
Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348–362.
Hankin, B. L. (2008). Stability of cognitive vulnerabilities to depression: A short-term prospective multiwave study. *Journal of Abnormal Psychology, 117*, 324–333.
Hankin, B. L., & Abramson, L. Y. (2001). Development of gender differences in depression: An elaborated cognitive vulnerability transactional stress theory. *Psychological Bulletin, 127*, 773–796.
Hankin, B. L., Lakdawalla, Z., Carter, I., Abela, J. R. Z., & Adams, P. (2007). Are neuroticism, cognitive vulnerability and self-esteem overlapping or distinct risks for depression? Evidence from exploratory and confirmatory factor analyses. *Journal of Social and Clinical Psychology, 26*, 29–63.

Heilman, K. M. (1997). The neurobiology of emotional experience. *Journal of Neuropsychiatry and Clinical Neuroscience, 9*, 439–448.

Heilman, R. M., Crisan, L. G., Houser, D., Miclea, M., & Miu, A. C. (2010). Emotion regulation and decision making under risk and uncertainty. *Emotion, 10*, 257–265.

Hoyle, R. H., Fejfar, M. C., & Miller, J. D. (2000). Personality and sexual risk taking; A quantitative review. *Journal of Personality, 68*, 1203–1231.

Indredavik, M. S., Vik, T., Evensen, K. A. I., Skranes, J., Taraldsen, G., & Brubakk, A. (2010). Perinatal risk and psychiatric outcome in adolescents born preterm with very low birth weight or term small for gestational age. *Journal of Developmental and Behavioral Pediatrics, 31*, 286–294.

Jacka, F. N., Kremer, P. J., Leslie, E. R., Berk, M., Patton, G. C., Toumbourou, J. W., et al. (2010). Associations between diet quality and depressed mood in adolescents: Results from the Australian Healthy Neighbourhoods Study. *Australian and New Zealand Journal of Psychiatry, 44*, 435–442.

Jacka, F. N., Pasco, J. A., Mykletun, A., Williams, L. J., Nicholson, G. C., Kotowicz, M. A., et al. (2011). Diet quality in bipolar disorder in a population-based sample of women. *Journal of Affective Disorders, 129*, 332–337.

Joo, M. (2010). Long-term effects of Head Start on academic and school outcomes of children in persistent poverty: Girls vs. boys. *Children and Youth Services Review, 32*, 807–814.

Katz, E. C., Fromme, K., & D'Amico, E. J. (2000). Effects of outcome expectancies and personality on young adults' illicit drug use, heavy drinking, and risky sexual behavior. *Cognitive Therapy and Research, 24*, 1–22.

Keating, D. (2004). Cognitive and brain development. In R. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 45–84). New York: Wiley.

Kilpatrick, D. G., Acierno, R., Saunders, B., Resnick, H. S., Best, C. L., & Schnurr, P. P. (2000). Risk factors for adolescent substance abuse and dependence: Data from a national sample. *Journal of Counseling and Clinical Psychology, 68*, 19–30.

Kim, J., & Cicchetti, D. (2010). Longitudinal pathways linking child abuse and neglect, emotion regulation, peer rejection, and psychopathology. *Journal of Child Psychology and Psychiatry, 51*, 706–716.

Kim, J., & Deater-Deckard, K. (2011). Dynamic changes in anger linking to developmental trajectories of internalizing and externalizing problems: The moderating role of attention. *Journal of Child Psychology and Psychiatry, 52*, 156–166.

Kim-Spoon, J., Ollendick, T., & Seligman, L. (2012). Perceived competence and depressive symptoms among adolescents: The moderating role of attributional style. *Child Psychiatry and Human Development, 43*, 612–630.

Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality, 71*, 1087–1112.

Kovacs, M., Joormann, J., & Gotlib, I. H. (2008). Emotion (dys)regulation and links to depressive disorders. *Child Development Perspectives, 2*, 149–155.

Krueger, R. F., Caspi, A., & Moffitt, T. E. (2000). Epidemiological personology: The unifying role personality in population-based research on problem behaviors. *Journal of Personality, 68*, 967–998.

Lane, R. D., & Schwartz, G. E. (1987). Levels of emotional awareness: A cognitive-developmental theory and its application to psychopathology. *American Journal of Psychiatry, 144*, 133–143.

Lautenbacher, S., Huber, C., Schöfer, D., Kunz, M., Parthum, A., Weber, P. G., et al. (2010). Attentional and emotional mechanisms related to pain as predictors of chronic postoperative pain: A comparison with other psychological and physiological predictors. *Pain, 151*, 722–731.

Lawson, K. R., & Ruff, H. A. (2004). Early attention and negative emotionality predict later cognitive and behavioural function. *International Journal of Behavioral Development, 28*, 157–165.

Lee, T. M. C., Chan, C. C. H., Leung, A. W. S., Fox, P. T., & Gao, J. (2009). Sex-related differences in neural activity during risk taking: An fMRI study. *Cerebral Cortex, 19*, 1303–1312.

Loesel, F., & Bliesener, T. (1994). Some high-risk adolescents do not develop conduct problems: A study of protective factors. *International Journal of Behavioral Development, 17*, 753–777.

Madden, G. J., & Bickel, W. K. (2009). *Impulsivity: The behavioral and neurological science of discounting*. Washington, DC: American Psychological Association.

Magar, E. C. E., Phillips, L. H., & Hoise, J. A. (2008). Self-regulation and risk-taking. *Personality and Individual Differences, 45*, 153–159.

Martin, L. N., & Delgado, M. R. (2011). The influence of emotion regulation on decisionmaking under risk. *Journal of Cognitive Neuroscience, 23*, 2569–2581.

Maughan, A., & Cicchetti, D. (2002). Impact of child maltreatment and interadult violence on children's emotion regulation abilities and socioemotional adjustment. *Child Development, 73*, 1525–1542.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *Journal of Neuroscience, 27*, 5796–5804.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*, 503–507.

McCord, J. (1990). Problem behaviors. In S. S. Feldman & G. Elliot (Eds.), *At the threshold: The developing adolescent* (pp. 414–430). Cambridge, MA: Harvard University Press.

McKelvey, L. M., Whiteside-Mansell, L., Bradley, R. H., Casey, P. H., Conners-Burrow, N. A., & Barrett, K. W. (2011). Growing up in violent communities: Do family conflict and gender moderate impacts on adolescents' psychosocial development? *Journal of Abnormal Child Psychology, 39*, 95–107.

McLeod, J. D., & Shanahan, M. J. (1996). Trajectories of poverty and children's mental health. *Journal of Health and Social Behavior, 37*, 207–220.

McRae, K., Ochsner, K. N., Mauss, I. B., Gabrieli, J. J. D., & Gross, J. J. (2008). Gender differences in emotion regulation: An fMRI study of cognitive reappraisal. *Group Processes & Intergroup Relations, 11*, 143–162.

Mitchell, D. K., & Murdock, K. K. (2005). Identifying risk and resource factors in children with asthma from urban settings: The context-health development model. *Journal of Asthma, 42*, 425–436.

Moilanen, K. L., Crockett, L. J., Raffaelli, M., & Jones, B. L. (2010). Trajectories of sexual risk from middle adolescence to early adulthood. *Journal of Research on Adolescence, 20*, 114–139.

Moore, M., Kirchner, L., Drotar, D., Johnson, N., Rosen, C., Ancoli-Israel, S., et al. (2009). Relationships among sleepiness, sleep time, and psychological functioning in adolescents. *Journal of Pediatric Psychology, 34*, 1175–1183.

Moore, M., Kirchner, H. L., Drotar, D., Johnson, N., Rosen, C., & Redline, S. (2011). Correlates of adolescent sleep time and variability in sleep time: The role of individual and health related characteristics. *Sleep Medicine, 12*, 239–245.

Nęcka, E. (2003). Intelligence and temperament. In R. J. Sternberg, J. Lautrey, & T. I. Lubart (Eds.), *Models of intelligence: International perspectives* (pp. 293–309). Washington DC: American Psychological Association.

Oddy, W. H., Robinson, M., Ambrosini, G. L., O'Sullivan, T. A., de Klerk, N. H., Beilin, L. J., et al. (2009). The association between dietary patterns and mental health in early adolescence. *Preventative Medicine, 49*, 39–44.

Odum, A. L., Madden, G. J., Badger, G. J., & Bickel, W. K. (2000). Needle sharing in opioid-dependent outpatients: Psychological processes underlying risk. *Drug and Alcohol Dependence, 60*, 259–266.

Oldehinkel, A. J., Hartman, C. A., Ferdinand, R. F., Verhulst, F. C., & Ormel, J. (2007). Effortful control as modifier of the association between negative emotionality and adolescents' mental health problems. *Development and Psychopathology, 19*, 523–539.

Olson, S. L., Sameroff, A. J., Kerr, D. C. R., Lopez, N. L., & Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: The role of effortful control. *Development and Psychopathology, 17*, 25–45.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping, 25*, 46–59.

Patterson, C. J. (2008). *Child development*. New York: McGraw-Hill Companies.

Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences, 9*, 60–68.

Piaget, J. (2000). Piaget's theory. In K. Lee (Ed.), *Childhood cognitive development: The essential readings* (pp. 33–47). Maiden, MA: Blackwell.

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology, 12*, 427–441.

Posner, M. I., & Rothbart, M. K. (2005). Influencing brain networks: implications for education. *Trends in Cognitive Sciences, 9*, 99–103.

Posner, M. I., & Rothbart, M. K. (2006). *Educating the human brain*. Washington DC: APA Books.

Quay, H. C. (1993). The psychobiology of undersocialized aggressive conduct disorder. *Development and Psychopathology, 5*, 165–180.

Resnick, M., Bearman, P., Blum, R., Bauman, K., Harris, K., Jones, J., et al. (1997). Protecting adolescents from harm: Findings from the National Longitudinal study on Adolescent Health. *Journal of the American Medical Association, 278*, 823–832.

Rojas, Y., & Stenberg, S. (2010). Early life circumstances and male suicide: A 30-year follow-up of a Stockholm cohort born in 1953. *Social Science and Medicine, 70*, 420–427.

Romer, D., Black, M., Ricardo, I., Feigelman, S., Kaljee, L., Galbraith, J., et al. (1994). Social influences on the sexual behavior of youth at risk for HIV exposure. *American Journal of Public Health, 84*, 977–985.

Rosenblitt, J. C., Soler, H., Johnson, S. E., & Quadagno, D. M. (2001). Sensation seeking and hormones in men and women: Exploring the link. *Hormones and Behavior, 40*, 396–402.

Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology* (Vol. 3). Social, emotional and personality development (5th ed., pp. 105–176). New York: Wiley.

Rothbart, M. K., & Bates, J. E. (2006). Temperament. In W. Damon, R Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology* (Vol. 3). Social, emotional, and personality development (6th ed., pp 99–166). New York: Wiley.

Rothbart, M. K., Ellis, L. K., Rueda, M. R., & Posner, M. I. (2003). Developing mechanisms of temperamental effortful control. *Journal of Personality, 71*, 1113–1143.

Rydell, A., Berlin, L., & Bohlin, G. (2003). Emotionality, emotion regulation, and adaptation among 5- to 8-year-old children. *Emotion, 3*, 30–47.

SAMHSA. (2005). Results from the 2004 National Survey on Drug Use and Mental Health. DHHS Publication no. SMA 05-4062 NSDUH Series H-28. Rockville, MD: Author.

Sanson, A., Hemphill, S. A., & Smart, D. (2004). Connections between temperament and social development: A review. *Social Development, 13*, 142–170.

Santelli, J. S., Lowry, R., Brener, N. D., & Robin, L. (2000). The association of sexual behaviors with socioeconomic status, family structure, and race/ethnicity among US adolescents. *American Journal of Public Health, 90*, 1582–1588.

Schachner, D. A., & Shaver, P. R. (2004). Attachment dimensions and sexual motives. *Personal Relationships, 11*, 179–195.

Selner-O'Hagan, M. B., Kindlon, D. J., Buka, S. L., Raudenbush, S. W., & Earls, F. J. (1998). Assessing exposure to violence in urban youth. *Journal of Child Psychology and Psychiatry, 39*, 215–224.

Shamosh, N. A., DeYoung, C. G., Green, A., Reis, D., Johnson, M. R., Conway, A. R. A., et al. (2008). Individual differences in delay discounting: Relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science, 19*, 904–911.

Silk, J. S., Steinberg, L., & Morris, A. S. (2003). Adolescents' emotion regulation in daily life: Links to depressive symptoms and problem behavior. *Child Development, 74*, 1869–1880.

Smith, C. A., Elwyn, L. J., Ireland, T. O., & Thornberry, T. P. (2010). Impact of adolescent exposure to intimate partner violence on substance use in early adulthood. *Journal of Studies on Alcohol and Drugs, 71*, 219–230.

Stein, G. L., Curry, J. F., Hersh, J., Breland-Noble, A., Silva, S. G., Reinecke, M. A., et al. (2010). Ethnic differences at the initiation of treatment for adolescent depression. *Cultural Diversity and Ethnic Minority Psychology, 16*, 152–158.

Steinberg, L. (2002). *Adolescence*. New York: McGraw-Hill.

Steinberg, L. (2004). Risk-taking in adolescence: What changes, and why? *Annals of the New York Academy of Sciences, 1021*, 51–58.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*, 78–106.

Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Developmental Psychobiology, 52*, 216–224.

Strelau, J., Zawadzki, B., & Piotrowska, A. (2001). Temperament and intelligence: A psychometric approach to the links between both phenomena. In J. M. Collis & S. Messick (Eds.), *Intelligence and personality: Bridging the gap in theory and measurement* (pp. 61–78). Mahwah, NJ: Erlbaum.

Thayer, R. E., Newman, J. R., & McClain, T. M. (1994). Self-regulation of mood: Strategies for changing a bad mood, raising energy, and reducing tension. *Journal of Personality and Social Psychology, 67*, 910–925.

Trute, B., Benzies, K. M., Worthington, C., Reddon, J. R., & Moore, M. (2010). Accentuate the positive to mitigate the negative: Mother psychological coping resources and family adjustment in childhood disability. *Journal of Intellectual and Developmental Disability, 35*, 36–43.

Tubman, J., Windle, M., & Windle, R. C. (1996). Cumulative sexual intercourse patterns among middle adolescents: Problem behavior precursors and concurrent health risk behaviors. *Journal of Adolescent Health, 18*, 182–191.

Van De Ven, M. O. M., Greenwood, P. A., Engels, R. C. M. E., Olsson, C. A., & Patton, G. C. (2010). Patterns of adolescent smoking and later nicotine dependence in young adults: A 10-year prospective study. *Public Health, 124*, 65–70.

van der Hulst, M., Vollenbroek-Hutten, M. M., Schreurs, K. M., Rietman, J. S., & Hermens, H. J. (2010). Relationships between coping strategies and lumbar muscle activity in subjects with chronic low back pain. *European Journal of Pain, 14*, 640–647.

Wahlstrom, D., White, T., & Luciana, M. (2010). Neurobehavioral evidence for changes in dopamine system activity during adolescence. *Neuroscience & Biobehavioral Review, 34*, 631–648.

Wallace, J. M., Brown, T. N., Bachman, J. G., & Laveist, T. A. (2003). The influence of race and religion on abstinence from alcohol, cigarettes and marijuana among adolescents. *Journal of Studies on Alcohol, 64*, 843–848.

Wardle, M. C., Gonzalez, R., Bechara, A., & Martin-Thormeyer, E. M. (2011). Iowa Gambling Task performance and emotional distress interact to predict risky sexual behavior in individuals with dual substance and HIV diagnoses. *Journal of Clinical and Experimental Neuropsychology, 32*, 1110–1121.

Webb, S., Monk, C., & Nelson, C. (2001). Mechanisms of postnatal neurobiological development: Implications for human development. *Developmental Neuropsychology, 19*, 147–171.

Widome, R., Neumark-Sztainer, D., Hannan, P. J., Haines, J., & Story, M. (2009). Eating when there is not enough to eat: Eating behaviors and perceptions of food among food-insecure youths. *American Journal of Public Health, 99*, 822–828.

Wilkosz, M. E., Chen, J.-L., Kennedy, C., & Rankin, S. (2010). Gender and ethnic disparities contributing to overweight in California adolescents. *Journal of Public Health, 18*, 131–144.

Williams, P. G., Holmbeck, G. N., & Greenley, R. N. (2002). Adolescent health psychology. *Journal of Consulting and Clinical Psychology, 70*, 828–842.

Wilson, M., & Daly, M. (2004). Do pretty women inspire men to discount the future? *Proceedings of the Royal Society B: Biological Sciences, 7*(271, Suppl 4), S177–S179.

Yen, C., King, B. H., & Tang, T. (2010). The association between short and long nocturnal sleep durations and risky behaviours and the moderating factors in Taiwanese adolescents. *Psychiatry Research, 179*, 69–74.

Zuckerman, M. (1983). *Biological bases of sensation seeking, impulsivity, and anxiety*. Hillsdale, NJ: Erlbaum.

Zuckerman, M. (1994). Impulsive unsocialized sensation seeking: The biological foundations of a basic dimension in personality. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 219–255). Washington, DC: American Psychological Association.

Chapter 5
Stressors and Vulnerabilities in Middle and Old Age: Opportunities for Prevention

George W. Rebok, Jeanine M. Parisi, and Alexandra M. Kueider

Introduction

The United States is currently in the middle of a profound demographic shift as the population rapidly ages and Americans are living longer. Within 20 years, older adults will make up almost one-quarter of the population. With recent advances in the medical field and decreases in age-specific mortality, life expectancy has increased dramatically over the last century from approximately 49 years of age at the turn of the twentieth century to almost 78 years of age today (U.S. Census Bureau, 2011), with women living almost 7 years longer than men (Frankenberg & Thomas, 2011). Depending on location, birth cohort, and other individual characteristics, the period of old age could extend over 40 years for some individuals (e.g., from ages 55–95).

Today, older adults make up a larger segment of the population than at any other time in history. The US Census Bureau estimates that this aging trend will continue to be affected by the baby boomer generation until 2050 (U.S. Census Bureau, 2011). With age come the knowledge, experience, and wisdom of growing older. However, aging also comes with an increase in the prevalence of disease and disability, which will undoubtedly impact social, economic, and political policies as the country grapples with how to care for an aging population. Furthermore, reducing the likelihood of disease occurrence would result in substantial benefits through improved quality of life, prolonged independence, and reduced cost to the healthcare system (Larson et al., 2006). Although there are several positive aspects of aging, the focus of this chapter is on age-related physical, cognitive, and psychological stressors and vulnerabilities. With this said, sometimes even the most vulnerable individuals seem to fare well in later adulthood. Successful aging is impacted by an individual's ability to adapt to physical, cognitive, emotional, and

G.W. Rebok (✉)
Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway Ave., Baltimore, MD 21205, USA
e-mail: grebok@jhsph.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_5,

Table 5.1 Factors influencing successful aging

Physical	Cognitive	Emotional	Behavioral
Cancers	AD	Anxiety	Alcohol use
Cardiovascular disease	MCI	Depression	Tobacco use
Diabetes	Vascular dementia	Emotional intelligence	Substance use
Hyperlipidemia	Cognitive reserve	Emotional support	Exercise/physical activity
Hypertension	Education	Personal control	Diet/nutrition
Frailty	Resiliency	Self-efficacy	Cognitive engagement
Obesity		Self-esteem	Social engagement
Osteoporosis		Stress	Coping mechanisms
Sarcopenia			Wisdom
Sensory impairments			
Stroke			

behavioral factors that can influence health and well-being through active engagement and changed expectations and goals (Baltes & Baltes, 1990). It includes individuals with a low risk of disease and disability and high cognitive and physical capacity (Rowe & Kahn, 1997). Therefore, the identification of modifiable factors and opportunities for prevention that may protect against age-related vulnerabilities is of great interest and importance to an aging population (Table 5.1).

Physical, Cognitive, and Emotional Vulnerabilities

Aging is a complex phenomenon that cannot be explained by a single mechanism that controls both the physiological and psychological aspects of aging. Visible signs of aging begin to appear around middle adulthood and result from a culmination of physical, cognitive, and emotional vulnerabilities that are affected by environmental and genetic factors. Although major medical advances have increased life expectancy dramatically over the last century, older adults still face a substantial number of cognitive and physical challenges related to aging (e.g., memory decline, frailty, sensory impairments), as well as chronic diseases like diabetes and cardiovascular-related diseases. Ethnic minorities (i.e., African Americans, Native Americans) have greater mortality and morbidity rates at every point in the life course compared with European Americans. Higher rates of poverty, lower socioeconomic status, biological dispositions, and stress stemming from inequality are among the factors contributing to higher rates of mortality and morbidity in minority populations (Jackson, Govia, & Sellers, 2011).

Physical Vulnerabilities

A host of physiological changes typically occur as one ages. For example, there are marked changes in cardiac performance, lung capacity, muscle mass, and sensory

functioning with age. For some individuals, these changes may have relatively little impact on everyday functioning, whereas for others these declines can lead to impaired physical functioning or could be indicative of more severe diseases. Moreover, chronic diseases tend to be highly correlated with disability and poor quality of life (Motl & McAuley, 2010), thus further complicating the aging process.

The most prevalent physical conditions among older adults are losses or impairments in hearing and/or vision. These sensory impairments have a significant impact on physical and psychological functioning and are a serious health problem in older populations (Brown & Barrett, 2011; Hildalgo et al., 2008; Weber & Wong, 2010). Additionally, approximately 5–20 % of adults older than age 70 are diagnosed with dual sensory loss; these estimates will include between 3.5 and 14 million baby boomers by 2030 (Brennan & Bally, 2007). Functional limitations and disabilities also are common in geriatric populations and encompass a broad spectrum of disorders that prevent an individual from completing a designated task. Frailty, described as a physical or functional decline that can occur in the presence or absence of pathology, is reported in 6–30 % of adults older than age 65 (Topinková, 2008; VanItallie, 2003) and affects 8.5 % of women and 4.1 % of men (Syddall et al., 2010). Frailty is associated with several negative outcomes, including faster physical and cognitive decline, disability, and increased risk of mortality (Fulop et al., 2010; Gill et al., 2002; VanItallie, 2003) and is often complicated by comorbid physical conditions including sarcopenia (degenerative loss of muscle mass and strength associated with aging) and osteopenia (loss of bone density).

Not all physical impairments are limited to the oldest ages; often they are present or begin to accumulate in midlife. Among adults ages 47 and older, cardiovascular diseases and diabetes mellitus (DM) accounted for 33 % and 3 % of deaths, respectively (Xu, Kockanek, Murphy, & Tejada-Vera, 2010). Vascular disorders such as DM, hypertension, hyperlipidemia, stroke, and cerebrovascular disease also have been identified as risk factors of dementia (Beeri, Ranova-Springer, Silverman, & Haroutunian, 2009; Kimm et al., 2010; Moroney et al., 1999; Panza et al., 2006; Rastas et al., 2010). Compared with European Americans, African Americans tend to have a higher incidence and prevalence of chronic health conditions (e.g., diabetes, hypertension, stroke, and vascular disease) placing them at a disproportionately greater risk for cognitive and physical impairments (Alzheimer's Association, 2011; Carson et al., 2011; Clark & Maddox, 1992; Clark, Mungai, Stump, & Wolinsky, 1997; Evans et al., 2003; Ford, Haug, Jones, Roy, & Folmar, 1990; Tang et al., 2001; Whitt-Glover, Taylor, Heath, & Macera, 2007). A population-based cohort study reported the incidence of hypertension was higher in African American participants than it was in European American participants (Carson et al., 2011). In addition, minority group status was found to be an independent risk factor for DM even after controlling for body mass index and socioeconomic status (Davidson, 2001). Risk for coronary heart disease (CHD) markedly increases with age, and among middle-aged adults, CHD is 2–5 times more common in men (Jousilahti, Vartiainen, Tuomilehto, & Puska, 1999). Some evidence shows that individuals who exhibit vascular disorders in midlife

(diagnosis before age 65) are more likely to have an increased risk of Alzheimer's disease (AD) and vascular dementia compared with individuals who developed vascular disorders after age 65 (Kimm et al., 2010), suggesting the importance of intervening at earlier ages for health and well-being later in life.

Cognitive Vulnerabilities

Along with these physical changes, there are also normative age-related changes in fluid cognitive abilities, such as processing speed, working memory capacity, inductive reasoning, spatial orientation, and word fluency (Schaie, 2005). While there are age-related declines in fluid abilities, there is also continued growth in other cognitive domains, such as language and knowledge over the life span (Baltes, 1997). Although it has been suggested that subtle cognitive changes occur as early as the mid-20s, most agree that mild cognitive changes can be noticed in the 40s and progress throughout midlife and beyond. However, there is considerable interindividual variability in the magnitude and rate of cognitive change.

Whereas some individuals remain relatively cognitively healthy well into old age, others develop more severe cognitive impairments, including dementia and AD. Mild cognitive impairment (MCI) is an intermediary stage between the normal forgetfulness associated with aging and the development of dementia. MCI is often associated with mild problems in thinking and memory that do not interfere with everyday functioning. Mildly cognitively impaired individuals with memory deficits and deficits in other cognitive domains are a particularly high-risk group for developing AD (Tabert et al., 2006). However, although many individuals with MCI will develop Alzheimer's disease, there is great variability depending on diagnostic criteria and sample selection.

AD affects approximately 5.4 million Americans. Although 5.2 million affected individuals are older than age 65, there are approximately 200,000 individuals with younger-onset AD (Alzheimer's Association, 2011). Briefly, AD is an advanced form of aging in the brain marked by the buildup of extracellular amyloid plaques and development of intracellular neurofibrillary tangles, which appear first in the entorhinal cortex of the limbic system, followed by the hippocampus, the temporal and parietal cortices, and finally the frontal lobes (Braak & Braak, 1991). Cognitive deficits parallel the structural decline and may begin with memory impairments; as the disease progresses, there are also marked changes in other cognitive abilities, as well as physical functioning, personality, and emotional affect. Today, AD is the fifth leading cause of death for individuals ages 65 and older in the United States, and it has been suggested that 16 million Americans will have the disease by 2050 without a way to prevent, cure, or slow its progression (Alzheimer's Association, 2011). Furthermore, epidemiological studies have reported higher prevalence rates of dementia (Gurland et al., 1999) and incidence rates of AD (Tang et al., 2001) in older African American adults compared with older European American adults.

Even in the absence of *APOE 4*, a gene associated with AD, African Americans are four times more likely to develop the disease than their European American counterparts (Tang et al., 2001). Research suggests that ethnic differences in rates of dementia are not attributed to differences in age, gender, or education, but rather seem to be driven by differences in cardiovascular factors and cognitive reserve (Ng, Leong, Chiam, & Kua, 2010).

Emotional Vulnerabilities

Individual psychological capacities and beliefs about one's life can shape and influence the ability to deal with age-associated cognitive and physical changes. Although research indicates that over time adults become more emotionally stable and agreeable (Charles & Horwitz, 2010), older adults face certain life challenges that can increase the risk of emotional vulnerabilities like chronic stress, depression, and anxiety.

Aging is characterized by a number of unique features (e.g., deteriorating physical health, loss of family and friends, exiting work roles) that can lead to chronic stress, which can influence the severity and type of disorders that an individual may develop over time. For example, chronic stress associated with caregiving can adversely impact multiple organ systems, including the immune and cardiovascular systems, leading to increased health risks for the caregiver (Roepke et al., 2011). Moreover, the accumulated stress associated with physical, cognitive, and emotional changes results in an integrated response of the neuroendocrine system, autonomic nervous system, and immune system, referred to as *allostasis*. Allostatic load, which can be thought of as the consequence of chronic exposure to daily stressors, may help explain individual differences in the aging process (Fig. 5.1). Many factors contribute to allostatic load including genes, early development, and lifestyle choices. The processes of allostatic load are not only associated with cardiovascular diseases and diabetes but are also associated with cognitive and emotional changes (McEwen, 2003).

Depression in later life is quite prevalent, with estimates ranging from 5 to 37 % for depressive disorders (Luppa et al., 2010) and could be potentially explained by high rates of medication use, physical disability, cognitive impairment, and socioeconomic challenges faced by this population. Women are at a significantly greater risk of depression compared with men, with depressive episodes occurring twice as frequently in women (American Psychiatric Association, 2000), a finding that holds true for European Americans, African Americans, and Latinos (George, 2011). Sex differences may be explained by hormonal changes, higher levels of stress, and lower levels of fulfillment in sex roles (Kessler, 2003). However, a large population-based study found that sex differences are only apparent until age 55, at which point the prevalence rates of depression markedly decline in women, which may result from hormonal changes related to menopause (Bebbington et al., 2003). In addition, research has suggested that stressful life experiences often

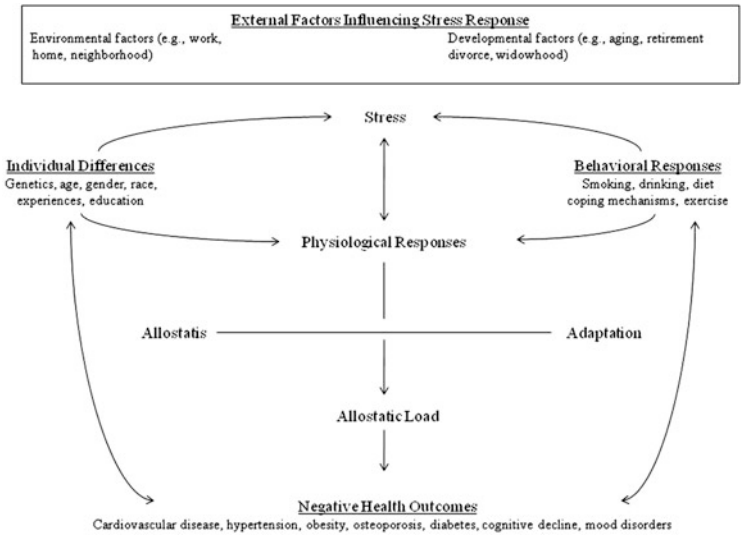

Fig. 5.1 Allostatic model of health – adapted from McEwen (1998)

precede the onset of depression and that the severity of this disorder may depend on the accumulation of multiple stressors (Lewinsohn, Hoberman, & Rosenbaum, 1988). The English Longitudinal Study of Aging reported that positive and negative exchanges in social relationships among middle-aged and older adults were independently associated with depression (Stafford, McMunn, Zaninottoa, & Nazroo, 2011).

Among older adults, anxiety is not as prevalent (ranging from 1 to 15 %) as it is in earlier adulthood, but it is still problematic (Bryant, Jackson, & Ames, 2008). Moreover, the comorbidity of depression and anxiety in the older population is quite high and reduces quality of life and functional status. For instance, within a large, community-based sample of adults between ages 55 and 85, almost 48 % of those with major depressive disorder also met criteria for anxiety disorders, whereas 26 % of those with anxiety disorders concurrently met criteria for major depressive disorder (Beekman et al., 2000). Clearly, prevention of these affective disorders is of great clinical significance for an older population that also faces age-related declines in cognitive and physical health.

Environmental Factors

Besides physiological processes, environmental factors contribute to age-associated changes in physical, cognitive, and emotional health. In particular, changes in the home and work environments, including changes in family structure

and dynamics (e.g., children move out of the family home; aging parents may move in) and retirement, may play a significant role in the health and well-being of middle-aged and older adults.

The Home Environment

Increased life expectancy, declining fertility, and increased participation of women in the workforce are among some of the many changes that alter family structure. Several changes occur within the home environment in middle and late adulthood. During middle adulthood significant transitions revolve around the family structure, including adolescent children leaving the home (e.g., empty nest syndrome), as well as adult children moving back into the family home. Additionally, if young children are still living at home, there may be a need for an older adult (grandparent) to move into the home to assist in the caretaking of the grandchildren.

Alternatively, in the face of cognitive and physical declines, a middle-aged adult may have to care for an aging parent in his or her home or consider seeking care in long-term facilities (e.g., assisted-living facilities, nursing homes). Research suggests that while taking care of an elderly parent can negatively impact the mental health of middle-aged adults, there is considerable variability in adaptation to this new source of stress. In fact, improvement in areas of subjective effectiveness and depressive symptoms was more common than deterioration (Townsend, Noelker, Deimling, & Bass, 1989). Moving from independent living to long-term care facilities is perhaps one of the most significant transitions an older adult will face. This stressful transition is often marked with uncertainty, stress, disorganization, and adaptation. Over time, however, older adults typically begin to grow comfortable with their new living situation and start building new relationships (Brandburg, 2007).

The Workplace

Employment provides a sense of fulfillment and financial stability and increases social interactions among older adults, but occupational choice may be constrained by personal or economic circumstances. In an economy where mass layoffs are the norm and unemployment rates are high, older adult may have to cope with unforeseen changes in occupational status that can impact economic stability and mental and physical well-being. Because of the interconnected nature of health and economic stability, changes in one area can have a significant impact on the other.

Retirement is a relatively new phenomenon. Historically, individuals continued to work until they were physically unable to do so. As medical improvements increased life expectancy, manual labor became less common in most jobs, and when the Social Security system was created, older adults were able to leave the

workforce and retire. Although some older adults embrace retirement and seek out new opportunities, others have difficulty with this stressful transition. Long-term distress levels in men increase after retirement, although they do report a short-term boost in morale shortly after retiring (Mandel & Roe, 2008).

Given the current economic recession, older adults are now deciding to delay retirement to secure better futures for themselves and their loved ones. Research has shown that financial strain has a significant impact on depressive symptoms—second only to the impact of physical illness (Krause, 1995; Krause & Baker, 1992). Furthermore, although working longer, older adults may face age-related discrimination in the workplace despite legislation enacted in the late 1960s that prevents employers from discriminating against persons older than age 40 (U.S. Department of Labor, 2007), thus contributing to work-related stress. Research has also suggested that work-related stress may potentially impact cognitive and physical functioning and accelerate the aging process (Piazza, Almeida, Dmitrieva, & Klein, 2010).

Behavioral Responses

Although both individual and environmental factors contribute to the aging process, there is great variability in the rate of age-associated cognitive and physical change. Part of this variability may be explained by an individual's behavioral response to daily life experiences. As such, employing strategies to successfully cope with life demands, having positive social support networks, and maintaining a healthful lifestyle can promote well-being in later adulthood.

Resiliency

As older adults age, they can encounter great difficulty in coping with age-related stress, which is often associated with declines in health and financial resources, as well as reduced autonomy (O'Hara et al., 2009). As previously noted, chronic stress can be disruptive to an individual's physiological and psychological homeostasis (Ingram & Luxton, 2005) and increase the likelihood of developing impairments in physical, cognitive, and emotional functioning. Despite inherent difficulties associated with aging, not all older adults seem to be vulnerable to stress, and some appear resilient to its negative effects. *Resiliency* can be thought of as one's ability or capacity to continue functioning normally following stress or loss. Research has demonstrated that individuals who are optimistic and socially engaged and have higher emotional well-being and fewer cognitive complaints, all characteristics associated with resiliency, are better able to positively adapt to stress (Lamond et al., 2008). The specific ways in which older adults respond to stressors, however, is influenced by the individual's level of vulnerability, the type of stress an

individual encounters, and biological, psychological, and social protective factors (Smyer & Qualls, 1999). These influences can either facilitate or buffer against stress or moderate its impact.

Personal Control

Feelings of personal control also may buffer against the negative effects of stress and contribute to improved cognitive, physical, and emotional health across the life course (Bandura, 1995; Krause, 2007; Lachman, 2006). Research has linked personal control to better physical and functional well-being, decreased health services use, and even longevity (Chipperfield, Campbell, & Perry, 2004; Chipperfield & Greenslade, 1999). One explanation is that individuals with a greater sense of control may frequently engage in health-promoting behaviors. Krause and Shaw (2000) found that individuals who reported a stronger sense of personal control were less likely to engage in unhealthy behaviors, such as smoking and drinking, as well as suffer from obesity, all of which may contribute to the development of other negative health conditions, including mortality.

Wisdom

Wisdom is often conceptualized as the pinnacle of successful aging (Baltes & Staudinger, 2000), influencing older adults more than physical health, financial well-being, and the physical environment (Baltes, Smith, & Staudinger, 1992). Despite the aforementioned age-related declines, several reasons exist as to why wisdom may be associated with growing older. Through the accumulation of both positive and negative life experiences, there is continued growth in knowledge, social, emotional, and moral development (e.g., Baltes & Smith, 1990; Clayton & Birren, 1980; Erikson, 1982; Labouvie-Vief, 2003). In fact, older adults tend to be more skilled than younger adults in tasks involving social and practical problem-solving (e.g., life knowledge, coordination of life themes, future goals, life planning; Baltes et al., 1992; Labouvie-Vief, 2003) and tend to handle interpersonal conflicts more effectively (Baltes & Baltes, 1990; Labouvie-Vief, Hakim-Larson, DeVoe, & Schoeberlein, 1989). In turn, successful resolution of complex (and often stressful) situations may act as a catalyst for personal growth, potentially representing one pathway for the development of wisdom (Ardelt, 2000; Arlin, 1990). Wisdom also may be stimulated through the intergenerational transmission of knowledge to younger generations (Erikson, 1982; McAdams & de St. Aubin, 1992; Parisi et al., 2009). One such program, Experience Corps®, was designed to draw on, and potentially activate, the wisdom of older adults (Fried et al., 2004). In this community-based, intergenerational program, older adults are trained and placed into elementary school classrooms to assist with academic goals and

behavioral management of school children (Rebok et al., 2011). Sustained participation in meaningful, generative roles may be conducive to the accumulation and refinement of wisdom through self-reflection and engagement with environments that promote the expansion of care beyond oneself toward others (Erikson, 1982; Parisi et al., 2009).

Lifestyle Factors

Although genetic and environmental factors play a large role in the development of disease, lifestyle factors can modify their effects. Individuals may protect themselves against age-related impairments by modifying their lifestyle. As such, healthful lifestyle choices regarding education, diet, exercise, and intellectual and social activities can reduce the risk of developing physical, cognitive, and emotional disorders.

Education

Cognitive enrichment early in life may account for some variation in cognitive ability in adulthood. Numerous studies have found that greater educational attainment (as indicated by years of formal education) is associated with higher levels of cognitive performance (Albert et al., 1995; Schaie, 2005), as well as a lower risk for the development of AD (Bennett et al., 2003; Gatz et al., 2001) or dementia (Katzman, 1993). However, once symptoms of the disease are prevalent, the rate of cognitive decline appears to be accelerated in individuals with higher levels of education (Scarmeas & Stern, 2003), which may suggest that older adults begin to lose the capacity to draw on education to compensate for age-related declines that are biologically driven.

Occupation

Educational attainment is closely related to occupation, which may also help promote cognition and emotional well-being in adulthood. Research has suggested that performing substantively complex work (i.e., work that requires thought, independent judgment, multiple ill-defined decisions) may positively impact cognitive performance (Schooler, Mulatu, & Oates, 1999). Additionally, longitudinal data from the Maastricht Longitudinal Study showed that 3 years after a baseline measurement at which no participants (ages 50–80) showed cognitive impairment, 1.5 % of those with mentally demanding jobs developed cognitive impairment, while among those with relatively light job demands, 4 % of the sample developed some cognitive impairment (Bosma et al., 2003). As such individuals who consistently place significant demands on their intellectual resources (i.e., multiple and

complex decisions, ill-defined problem-solving) may maintain or even enhance cognitive potential (e.g., Schaie, 2005; Schooler & Mulatu, 2001). Furthermore, the longer an individual is exposed to a complex environment, the greater the absorption of knowledge, skills, and competencies this exposure affords.

Cognitive and Social Engagement

Consistent with the notion of environmental complexity (Schooler, 1987), several studies suggest that a lifestyle characterized by engagement in intellectual and social leisure activities is associated with slower cognitive decline, reduced risk of dementia, and even decreased risk of mortality (e.g., Bygren, Konlaan, & Johansson, 1996; Crowe, Andel, Pedersen, Johansson, & Gatz, 2003; Holtzman et al., 2004; Hultsch, Small, Hertzog, & Dixon, 1999; Konlaan, Theobald, & Bygren, 2002; Scarmeas, Levy, Tang, Manly, & Stern, 2001; Seeman, Lusignolo, Albert, & Berkman, 2001; Verghese et al., 2003; Wang, Karp, Winblad, & Fratiglioni, 2002; Wilson & Bennett, 2003). One possibility is that maintaining an active lifestyle may preserve skills or repertoires that allow an individual to cope for a longer period before the clinical manifestations of dementia emerge (e.g., cognitive reserve; Scarmeas et al., 2001; Wang et al., 2002).

Exercise

Accumulating evidence suggests that physical activity is associated with several positive outcomes in later life, including improvements in cognitive performance, as well as a decreased risk of AD and dementia (e.g., Albert et al., 1995; Clarkson-Smith & Hartley, 1989; Kramer et al., 1999; Larson et al., 2006; Laurin, Verreault, Lindsay, MacPherson, & Rockwood, 2001; Podewils et al., 2005). A meta-analysis by Colcombe and Kramer (2003) showed that reduced loss of hippocampal brain tissue in the aging brain is related to the level of physical fitness. Similarly, a systematic review of the relationship between physical activity and functional limitation, disability, and loss of independence found that greater aerobic physical activity was associated with higher functional status and independence (Paterson & Warburton, 2010). Collectively, these findings indicate that the protective benefits from exercise may have a general effect on slowing the progression of AD and may be of particular importance to individuals who are at high risk for cognitive and physical decline.

Diet and Nutrition

Diet may further aid in the promotion of health and well-being in later adulthood. As certain nutrients are essential for brain function, studies have suggested the potential role of omega-3 fatty acids, antioxidants, vitamin D, and B vitamins as a

means of slowing cognitive decline or decreasing risk for dementias and disease (Gillette Guyonnet et al., 2007; Rutten, Steinbusch, Korr, & Schmitz, 2002; Shepherd, 2009). Unfortunately, research findings suggest that many older adults are not getting the recommended daily allowance of some nutrients. The USDA Human Nutrition Research Center on Aging found deficiencies in calcium; zinc; iron; magnesium; vitamins B6, B12, D, and E; and folic acid (Bliss, 2007). Besides the potential impact on cognitive functioning, current research suggests that nutrition plays an important role in the prevention and treatment of many age-related chronic diseases (Shepherd, 2009).

Substance Use

Physiological changes associated with aging increase sensitivity to drugs creating the potential for neurotoxicity and adverse consequences. Not all drug use is harmful, and although results are mixed, some suggest that moderate amounts of alcohol are protective (Gross et al., 2011; Stampfer, Kang, Chen, Cherry, & Grodstein, 2005). Although heavy alcohol consumption impairs the brain and is associated with an increased risk of dementia, moderate alcohol consumption in women has been associated with better cognitive function compared with non-drinkers, suggesting one drink per day does not appear to impair cognitive function. The protective effect of moderate alcohol use may be a function of the social environment in which alcohol is typically consumed or due to the decreased risk of cardiovascular disease in moderate drinkers.

A common misperception, however, is that substance abuse and dependence occur only in younger adults. Current drug use among older adults has surpassed previously predicted estimates and indicates that healthcare providers need to increase efforts to screen adults of all ages for substance abuse. Prevalence rates of substance use disorders in older adults are projected to more than double by 2020 (Substance Abuse and Mental Health Services Administration, 2009), resulting in a substantial public health burden. Currently, most substance abuse treatment focuses on younger patients, but as the baby boomer generation continues to age, an integrated approach is necessary to address its growing needs. Interventions directly targeting older adults have been successful in increasing the number of older substance users identified and treated, as well as improving substance use behaviors in the population (Schonfeld et al., 2010).

Almost 90 % of older adults use prescription and over-the-counter medications to treat various physical and mental health ailments. Medication misuse is especially an issue because older adults generally use multiple medications and are more prone to errors in following prescription instructions. In a large sample of almost 3,500, older adults' prescription medication misuse was the most common substance use problem, followed by alcohol abuse, over-the-counter medication misuse, and illicit substances (Schonfeld et al., 2010).

Conclusion and Future Directions

The number of individuals ages 60 and older in the United States will increase from 19 % currently to approximately 40 % by 2050, placing a greater percentage of Americans at risk for physical, cognitive, and emotional vulnerabilities. Therefore, the search for effective prevention strategies that will delay onset of and slow decline in physical and cognitive functioning and prevent the occurrence of later life mental and emotional disorders is becoming increasingly important.

Interventions targeting underlying physical impairments have shown promise in reducing the progression of functional decline in older adults (e.g., Gill et al., 2002; VanItallie, 2003; Waters, Baumgartner, Garry, & Vallas, 2010). Therefore, simple interventions, such as increasing physical activity may prove critical in delaying the onset of cognitive and functional decline (Tan, Xue, Li, Carlson, & Fried, 2006), especially for individuals who are more susceptible to the development of dementia. Although it is reasonable to expect a relationship between mentally stimulating activities and cognition, it may be less clear why there should exist a relationship between physical activity and cognition. One explanation can be found in the well-documented relationship between physical activity and disease. A large literature now exists showing the health benefits of physical activity and exercise for the primary and secondary prevention of cardiovascular disease, type 2 diabetes, colon and breast cancer, and osteoporosis (Bassuk & Manson, 2005; Kruk, 2007). It has also been reported that cardiovascular disease, diabetes, and cancer are associated with impaired cognition (Bruce et al., 2008; Ek, Almkvist, Wiberg, Stragliotto, & Smits, 2010; Knecht, Wersching, Lohmann, Berger, & Ringelstein, 2009; Knopman et al., 2009). Therefore, preventive interventions aimed at improving physical activity may be expected to improve cognition through reduction of risk of diseases associated with cognitive decline (Hertzog, Kramer, Wilson, & Linderberger, 2009).

Likewise, randomized, controlled trials, such as the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) trial, show promising results for the long-term benefits of cognitive interventions, potentially delaying cognitive and functional decline later in life (e.g., Ball et al., 2002; Rebok et al., 2014; Willis et al., 2006). The ACTIVE trial was designed to compare the effects of training on three different cognitive abilities (episodic memory, inductive reasoning, speed of information processing) showing significant age-related declines and thought to underlie the performance of everyday functional tasks (e.g., food preparation, driving, medication use, financial management) in a large ($N = 2{,}832$) and ethnically diverse (26 % African American) sample recruited from six metropolitan areas. The trial phase consisted of baseline assessment followed by random assignment to one of the cognitive interventions, immediate posttest, booster training for a subsample, and posttests at 1, 2, 3, 5 and 10 years. The interventions were standardized to consist of ten sessions lasting 60–75 min completed over 5–6 weeks. In the short term, each ACTIVE intervention produced immediate, reliable improvements on the cognitive ability targeted for training. These improvements remained statistically significant at 2-, 3-, and 5-years, and for reasoning and speed at the 10-year follow-up, despite some

dissipation of the training effects over time. There was also evidence that more intensive training in the form of a booster resulted in better long-term maintenance of the cognitive improvements for the reasoning and speed of processing interventions. Most importantly, the training generalized to the primary outcome of daily function over the long term (Willis et al., 2006). All three ACTIVE intervention groups produced comparably sized positive effects on self-reported activities of daily living at 10 years after training. The results from ACTIVE and other randomized trials (Smith et al., 2009) provide growing support for behavioral interventions designed to prevent cognitive declines and functional losses with age.

With regard to emotional vulnerabilities, interventions for minimizing the negative impact of stress or training appropriate coping mechanisms may help older adults face daily life stressors, as well as more situational challenges (e.g., dealing with loss of a loved one). This may become clinically important in situations where older adults' emotional vulnerabilities have led to the development of a depressive disorder and they are actively undergoing treatment. For example, efforts to increase perceptions of control through cognitive behavioral therapy may enhance the outcome of pharmacological treatments for depression in older adults and prevent relapse after therapy (de Beurs et al., 2005).

In the absence of dementia, older adults seem to cope as effectively as younger adults and may even be more efficient in their use of various coping and mastery strategies (Aldwin, 2007). Having endured major stressors and traumas earlier in life, older adults may have developed effective coping styles, making them better equipped to handle everyday life problems. Under such conditions, both positive and tolerable stress can contribute to resilience. In fact, such circumstances may give rise to stress inoculation, through which an individual becomes immune to particular stressors after having tolerated the effects of other similar stressors. People learn to deal with stress throughout their life spans, and the degree that they can successfully cope may have more to do with how they appraise the stressors than with the amount of stress exposure per se (Aldwin, Sutton, Chiara, & Spiro, 1996). Those who can see the positive aspects of even the most dire situations are less likely to suffer from the negative outcomes of stress such as post-traumatic stress disorder. However, older adults' ability to cope with life's challenges may become severely compromised when faced with serious cognitive losses, underscoring the importance of interventions to prevent later life cognitive decline and dementia.

Although this chapter focused on mid- and late-life environmental risk factors, it should be noted that genetic and early-life factors are equally important for shaping cognitive, physical, and mental health in later adulthood. We need to consider how to promote health and well-being at younger ages, hence the need for a life-course perspective. Additionally, given that the occurrence of dementia later in life may be due to a cumulated risk during the whole life span, research is needed to determine whether a critical period (e.g., early in life, midlife, later adulthood) exists in which the effects of intervention have the greatest impact on cognition and psychological and emotional well-being. By targeting modifiable risk behaviors earlier in the life course, we may be able to reduce burden of disease and disability in later life.

References

Albert, M. S., Jones, K., Savage, K. R., Berkman, L., Seeman, T., Blazer, D., et al. (1995). Predictors of cognitive change in older persons: MacArthur studies of successful aging. *Psychology and Aging, 10*, 578–589.

Aldwin, C. M. (2007). *Stress, coping, and development: An integrative approach*. New York, NY: Guilford.

Aldwin, C. M., Sutton, K. J., Chiara, G., & Spiro, A. (1996). Age differences in stress, coping, and appraisal: Findings from the Normative Aging Study. *The Journals of Gerontology, 51B*, 179–188.

Alzheimer's Association. (2011). Alzheimer's facts and figures. Retrieved from http://www.alz.org/alzheimers_disease_facts_and_figures.asp

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (Rev. ed.)* (4th ed.). Washington, DC: Author.

Ardelt, M. (2000). Intellectual versus wisdom-related knowledge: The case for a different kind of learning in the later years of life. *Educational Gerontology, 26*, 771–789.

Arlin, P. K. (1990). Wisdom: The art of problem finding. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development*. New York, NY: Cambridge University Press.

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., et al. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of American Medical Association, 288*, 2271–2281.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist, 52*, 366–380.

Baltes, P. B., & Baltes, M. M. (1990). Selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences*. New York, NY: Cambridge University Press.

Baltes, P. B., & Smith, J. (1990). Toward a psychology of wisdom and its ontogenesis. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development*. New York, NY: Cambridge University Press.

Baltes, P. B., Smith, J., & Staudinger, U. M. (1992). Wisdom and successful aging. In T. B. Sonderegger (Ed.), *Nebraska symposium on motivation*. Lincoln, NB: University of Nebraska Press.

Baltes, P. B., & Staudinger, U. M. (2000). Wisdom. A metaheuristic to orchestrate mind and virtue toward excellence. *American Psychologist, 55*, 122–136.

Bandura, A. (1995). *Self-efficacy in changing societies*. New York, NY: Cambridge University Press.

Bassuk, S. S., & Manson, J. E. (2005). Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. *Journal of Applied Physiology, 99*, 1193–1204.

Bebbington, P., Dunn, G., Jenkins, R., Lewis, G., Brugha, T., Farrell, M., et al. (2003). The influence of age and sex on the prevalence of depressive symptoms: Report from the National Survey on Psychiatric Morbidity. *International Review of Psychiatry, 15*, 74–83.

Beekman, A. T., de Beurs, E., van Balkom, A. J. L. M., Deeg, D. J. H., van Dyck, R., & van Tilburg, W. (2000). Anxiety and depression in later life: Co-occurrence and communality of risk factors. *American Journal of Psychiatry, 157*, 89–95.

Beeri, M. S., Ranova-Springer, R., Silverman, J. M., & Haroutunian, V. (2009). The effects of cardiovascular risk factors on cognitive compromise. *Dialogues Clinical Neuroscience, 11*, 201–212.

Bennett, D. A., Wilson, R. S., Schneider, A., Evans, D. A., Mendes de Leon, C. F., Arnold, S. E., et al. (2003). Education modifies the relation of AD pathology to level of cognitive function in older persons. *Neurology, 60*, 1909–1915.

Bliss, R. M. (2007). Boosting our knowledge of brain food. *Agricultural Research, 55*, 14–16.

Bosma, H., van Boxtel, M., Ponds, R., Houx, P., Burdorf, A., & Jolles, J. (2003). Mental work demands protect against cognitive impairment: MAAS prospective cohort study. *Experimental Aging Research, 29*, 33–45.

Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer related changes. *Acta Neurophathology, 82*, 239–259.

Brandburg, G. (2007). Making the transition to nursing home life: A framework to help older adults adapt to the long-term care environment. *Journal of Gerontological Nursing, 33*, 50–56.

Brennan, M., & Bally, S. J. (2007). Psychosocial adaptations to dual sensory loss in middle and late adulthood. *Trends in Amplification, 11*, 281–300.

Brown, R. L., & Barrett, A. E. (2011). Visual impairment and quality of life among older adults: An examination of explanations for the relationship. *The Journals of Gerontology, 66B*(3), 364–373.

Bruce, D. G., Davis, W. A., Casey, G. P., Starkstein, S. E., Clarmette, R. M., Foster, J., et al. (2008). Predictors of cognitive impairment and dementia in older people with diabetes. *Diabetologia, 51*, 241–248.

Bryant, C., Jackson, H., & Ames, D. (2008). The prevalence of anxiety in older adults: Methodology and a review of the literature. *Journal of Affective Disorders, 109*, 233–250.

Bygren, L. O., Konlaan, B. B., & Johansson, S. E. (1996). Attendance at cultural events, reading books or periodicals, and making music or singing in a choir as determinants for survival: Swedish interview survey of living conditions. *British Medical Journal, 313*, 1577–1580.

Carson, A. P., Howard, G., Burke, G. L., Shea, S., Levitan, E. B., & Munter, P. (2011). Ethnic differences in hypertension incidence among middle-age and older adults. The multi-ethnic study of atherosclerosis. *Hypertension, 57*, 1101–1107.

Charles, S. T., & Horwitz, B. N. (2010). Positive emotions and health: What we know about aging. In C. A. Depp & D. V. Jeste (Eds.), *Successful cognitive and emotional aging* (pp. 157–172). Washington, DC: APA.

Chipperfield, J., Campbell, D., & Perry, R. (2004). Stability of perceived control: Implication of health among very old community-dwelling adults. *Journal of Aging and Health, 16*, 116–147.

Chipperfield, J. G., & Greenslade, L. (1999). Perceived control as a buffer in the use of health care services. *Journal of Gerontology: Psychological Sciences, 54*, 146–154.

Clark, D. O., & Maddox, G. L. (1992). Racial and social correlates of age-related changes in functioning. *Journal of Gerontology: Social Sciences, 47*, 222–232.

Clark, D. O., Mungai, S. M., Stump, T. E., & Wolinsky, F. D. (1997). Prevalence and impact of risk factors for lower body difficulty among Mexican Americans, African Americans, and Whites. *Journal of Gerontology: Medical Sciences, 52*, 97–105.

Clarkson-Smith, L., & Hartley, A. A. (1989). Relationships between physical exercise and cognitive abilities in older adults. *Psychology and Aging, 4*, 183–189.

Clayton, V. P., & Birren, J. E. (1980). The development of wisdom across the life-span: A reexamination of an ancient topic. *Life-Span Development and Behavior, 3*, 103–135.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14*, 125–130.

Crowe, M., Andel, R., Pedersen, N. L., Johansson, B., & Gatz, M. (2003). Does participation in leisure activities lead to reduced risk of Alzheimer's disease? A prospective study of Swedish twins. *Journal of Gerontology: Psychological Review, 58B*, 249–255.

Davidson, M. B. (2001). The disproportionate burden of diabetes in African-American and Hispanic populations. *Ethnicity and Disease, 11*, 148–151.

de Beurs, E., Cmijs, H., Twisk, J. W. R., Sonneberg, C., Beekman, A. T. F., & Deeg, D. (2005). Stability and change of emotional functioning in late life: Modeling of vulnerability profiles. *Journal of Affective Disorders, 84*, 53–62.

Ek, L., Almkvist, O., Wiberg, M. K., Stragliotto, G., & Smits, A. (2010). Early cognitive impairment in a subset of patients with presumed low-grade glioma. *Neurocase, 16*, 503–511.

Erikson, E. (1982). *The life cycle completed: A review*. New York, NY: Norton.

Evans, D. A., Bennett, D. A., Wilson, R. S., Bienias, J. L., Morris, M. C., Scherr, P. A., et al. (2003). Incidence of Alzheimer disease in a biracial urban community: Relation to apolipoprotein e allele status. *Archives of Neurology, 60*, 185–189.

Ford, A. B., Haug, M. R., Jones, P. K., Roy, A. W., & Folmar, S. J. (1990). Race-related differences among elderly urban residents: A cohort study, 1975–1984. *Journal of Gerontology: Social Sciences, 45*, 163–171.

Frankenberg, E., & Thomas, D. (2011). Global aging. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (7th ed., pp. 73–87). San Diego, CA: Elsevier Academic Press.

Fried, L. P., Carlson, M. C., Freedman, M., Frick, K. D., Glass, T. A., Hill, J., et al. (2004). A social model for health promotion for an aging population: Initial evidence on experience corps model. *Journal of Urban Health, 81*, 64–78.

Fulop, T., Larbi, A., Witkowski, J. M., McElhaney, J., Loeb, M., Mitnitski, A., et al. (2010). Aging, frailty and age-related diseases. *Biogerontology, 11*, 547–563.

Gatz, M., Svedberg, P., Petersen, N. L., Mortimer, J. A., Berg, S., & Johansson, B. (2001). Education and the risk of Alzheimer's disease: Findings from the study of dementia in Swedish twins. *Journal of Gerontology, 56*, 292–300.

George, L. K. (2011). Social factors, depression, and aging. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (7th ed., pp. 149–162). San Diego, CA: Elsevier Academic Press.

Gill, T. M., Baker, D. I., Gottschalk, M., Peduzzi, P. N., Allore, H., & Byers, A. (2002). A program to prevent functional decline in physically frail, elderly persons who live at home. *New England Journal of Medicine, 347*, 1068–1074.

Gillette Guyonnet, S., Abellan Van Kan, G., Andrieu, S., Barberger Gateau, P., Berr, C., Bonnefoy, M., et al. (2007). IANA task force on nutrition and cognitive decline with aging. *The Journal of Nutrition, Health & Aging, 11*, 132–152.

Gross, A. L., Rebok, G. W., Ford, D. E., Chu, A. Y., Gallo, J. J., Liang, K. Y., et al. (2011). Alcohol consumption and domain-specific cognitive function in older adults: Longitudinal data from the Johns Hopkins Precursors Study. *Journals of Gerontology Social Sciences, 66*, 39–47.

Gurland, B. J., Wilder, D. E., Lantigua, R., Stern, Y., Chen, J., Killeffer, E. H. P., et al. (1999). Rates of dementia in three ethnoracial groups. *International Journal of Geriatric Psychiatry, 14*, 481–493.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Linderberger, U. (2009). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest, 9*, 1–65.

Hildalgo, J. L.-T., Gras, C. B., Lapeira, J. M. T., Martínez, I. P., Verdejo, M. A. L., Rabadán, F. E., et al. (2008). The hearing-dependent daily activities scale to evaluate impact of hearing loss in older people. *Annals of Family Medicine, 6*, 441–447.

Holtzman, R. E., Rebok, G. W., Saczynski, J. S., Kouzis, A. C., Wilcox Doyle, K., & Eaton, W. W. (2004). Social network characteristics and cognition in middle-aged and older adults. *Journal of Gerontology: Psychological Sciences, 59B*, 278–284.

Hultsch, D. F., Small, B. J., Hertzog, C., & Dixon, R. A. (1999). Use it or lose it: Engaged lifestyle as a buffer of cognitive decline in aging? *Psychology and Aging, 14*, 245–263.

Ingram, R. E., & Luxton, D. (2005). Vulnerability-stress models. In B. L. Hankin & J. R. Z. Abela (Eds.), *Development of psychopathology: A vulnerability-stress perspective* (pp. 32–46). New York, NY: Sage.

Jackson, J. S., Govia, I. O., & Sellers, S. L. (2011). Racial and ethnic influences over the life course. In R. H. Binstock & L. K. George (Eds.), *Handbook of aging and the social sciences* (7th ed., pp. 91–103). San Diego, CA: Elsevier Academic Press.

Jousilahti, P., Vartiainen, E., Tuomilehto, J., & Puska, P. (1999). Sex, age, cardiovascular risk factors, and coronary heart disease: A prospective follow-up study of 14,786 middle-aged men and women in Finland. *Circulation, 99*, 1165–1172.

Katzman, R. (1993). Education and the prevalence of dementia and Alzheimer's disease. *Neurology, 43*, 13–20.

Kessler, R. C. (2003). Epidemiology and women and depression. *Journal of Affective Disorders, 74*, 5–13.

Kimm, H., Lee, P. H., Shin, Y. J., Park, K. S., Jo, J., Lee, Y., et al. (2010). Mid-life and late-life vascular risk factors and dementia in Korean men and women. *Archives of Gerontology and Geriatrics, 52*(3), e117–e122. doi:10.1016/j.archger.2010.09.004.

Knecht, S., Wersching, H., Lohmann, H., Berger, K., & Ringelstein, E. B. (2009). How much does hypertension affect cognition? Explained variance in cross-sectional analysis of non-demented community-dwelling individuals in the SEARCH study. *Journal of Neurological Science, 283*, 149–152.

Knopman, D. S., Roberts, R. O., Geda, Y. E., Boeye, B. F., Pankratz, V. S., Cha, R. H., et al. (2009). Association of prior stroke with cognitive function and cognitive impairment: A population-based study. *Archives of Neurology, 66*, 614–619.

Konlaan, B. B., Theobald, H., & Bygren, L. O. (2002). Leisure time activity as a determinant of survival: A 26-year follow-up of a Swedish cohort. *Public Health, 116*, 227–230.

Kramer, A. F., Hahn, S., Cohen, N. J., Banich, M. T., McAuley, E., Harrison, C. R., et al. (1999). Ageing, fitness and neurocognitive function. *Nature, 400*, 418–419.

Krause, N. (1995). Assessing stress-buffering effects: A cautionary note. *Psychology and Aging, 10*, 518–526.

Krause, N. (2007). Age and decline in role-specific feelings of control. *Journals of Gerontology: Social Sciences, 62*, 28–35.

Krause, N., & Baker, E. (1992). Financial strain, economic values, and somatic symptoms in later life. *Psychology and Aging, 7*, 4–14.

Krause, N., & Shaw, B. A. (2000). Role-specific feelings of control and mortality. *Psychology and Aging, 15*, 617–626.

Kruk, J. (2007). Physical activity in the prevention of the most frequent chronic diseases: An analysis of the recent evidence. *Physical Exercise and Chronic Diseases, 8*, 325–338.

Labouvie-Vief, G. (2003). Dynamic integration: Affect, cognition, and the self in adulthood. *Current Directions in Psychological Science, 12*, 201–206.

Labouvie-Vief, G., Hakim-Larson, J., DeVoe, M., & Schoeberlein, S. (1989). Emotions and self-regulation: A life span view. *Human Development, 32*, 279–299.

Lachman, M. E. (2006). Perceived control over aging-related declines: Adaptive beliefs and behaviors. *Current Directions in Psychological Science, 15*, 282–286.

Lamond, A. J., Depp, C., Allison, M., Langer, R., Reichstadt, J., Moore, D. J., et al. (2008). Measurement and predictors of resilience among community-dwelling older women. *Journal of Psychiatric Research, 43*, 148–154.

Larson, E. B., Wang, L., Bowen, J. D., McCormick, W. C., Teri, L., Crane, P., et al. (2006). Exercise is associated with reduced risk for incident dementia among persons 65 years of age and older. *Annals of Internal Medicine, 144*, 73–81.

Laurin, D., Verreault, R., Lindsay, J., MacPherson, K., & Rockwood, K. (2001). Physical activity and risk of cognitive impairment and dementia in elderly persons. *Archives of Neurology, 58*, 498–504.

Lewinsohn, P. M., Hoberman, H. M., & Rosenbaum, M. (1988). A prospective study of risk factors for unipolar depression. *Journal of Abnormal Psychology, 97*, 251–264.

Luppa, M., Sikorski, C., Luck, T., Ehreke, L., Konnopka, A., Wiese, B., et al. (2010). Age- and gender-specific prevalence of depression in latest life: Systematic review and meta-analysis. *Journal of Affective Disorders, 136*(3), 212–221. doi:10.1016/j.jad.2010.11.033.

Mandel, B., & Roe, B. (2008). Job loss, retirement and the mental health of older Americans. *Journal of Mental Health Policy and Economics, 11*, 167–176.

McAdams, D. P., & de St. Aubin, E. (1992). A theory of generativity and its assessment through self-report, behavioral acts, and narrative themes in autobiography. *Journal of Personality and Social Psychology, 62*, 1003–1015.

McEwen, B. S. (1998). Protective and damaging effects of stress mediators. *New England Journal of Medicine, 338*, 171–179.

McEwen, B. S. (2003). Interacting mediators of allostasis and allostatic load: Towards an understanding of resilience in aging. *Metabolism, 52*, 10–16. doi:10.1053/50026-049(03)00295-6.

Moroney, J. T., Tang, M., Berglund, L., Small, S., Merchant, C., Bell, K., et al. (1999). Low-density lipoprotein cholesterol and the risk of dementia with stroke. *Journal of American Medical Association, 281*, 254–260.

Motl, R. W., & McAuley, E. (2010). Physical activity, disability, and quality of life in older adults. *Physical Medicine and Rehabilitation Clinics of North America, 21*, 299–308. doi:10.1016/j.pmr.2009.12.006.

Ng, T. P., Leong, T., Chiam, P. C., & Kua, E. H. (2010). Ethnic variations in dementia: The contributions of cardiovascular, psychosocial and neuropsychological factors. *Dementia and Geriatric Cognitive Disorders, 29*, 131–138.

O'Hara, R., Beaudreau, S. A., Luzon, A., Hah, M., Hubbard, J. T., & Sommer, B. (2009). Stress, resilience, and the aging brain. In C. Depp & D. V. Jeste (Eds.), *Successful cognitive and emotional aging* (pp. 173–193). Washington, DC: American Psychiatric Publishing.

Panza, F., D'Introno, A., Colacicco, A. M., Capurso, C., Pichichero, G., Capurso, S. A., et al. (2006). Lipid metabolism in cognitive decline and dementia. *Brain Research Reviews, 51*, 275–292.

Parisi, J. M., Rebok, G. W., Carlson, M. C., Fried, L. P., Seeman, T. E., Tan, E. J., et al. (2009). Can the wisdom of aging be activated and make a difference societally? *Educational Gerontology, 35*, 867–879.

Paterson, D. H., & Warburton, D. E. R. (2010). Physical activity and functional limitations in older adults: A systematic review related to Canada's physical activities guidelines. *International Journal of Behavioral Nutrition and Physical Activity, 7*, 1–22.

Piazza, J. R., Almeida, D. M., Dmitrieva, N. O., & Klein, L. C. (2010). Frontiers in the use of biomarkers of health in research on stress and aging. *The Journals of Gerontology: Psychological Sciences and Social Sciences, 65B*, 513–525.

Podewils, L. J., Guallar, E., Kuller, L. H., Fried, L. P., Lopez, O. L., Carlson, M. C., et al. (2005). Physical activity, APOE genotype, and dementia risk: Findings from the Cardiovascular Health Cognition Study. *American Journal of Epidemiology, 161*, 639–651.

Rastas, S., Pirttila, T., Mattila, K., Verkkoniemi, A., Juva, K., Niinisto, L., et al. (2010). Vascular risk factors and dementia in the general population aged $>$ 85 years: Prospective population-based study. *Neurobiology of Aging, 31*, 1–7.

Rebok, G. W., Barron, J. S., Carlson, M. C., Frick, K. D., McGill, S., Parisi, J. M., et al. (2011). Experience Corps®: A civic engagement-based public health intervention. In P. E. Hartman-Stein & A. La Rue (Eds.), *Enhancing cognitive fitness in adults: A handbook for the development of community based programs* (pp. 469–488). New York, NY: Springer.

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., et al. (2014). Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. *Journal of the American Geriatric Society, 62*, 16–24.

Roepke, S. K., Allison, M., von Känel, R., Mausbach, B. T., Chattillion, E. A., Harmell, A. L., et al. (2011). Relationship between chronic stress and carotid intima-media thickness (IMT) in elderly Alzheimer's caregivers. *Stress, 15*(2), 121–129.

Rowe, J. W., & Kahn, R. L. (1997). Successful aging. *Gerontologist, 37*, 433–440.

Rutten, B. P. F., Steinbusch, H. W. M., Korr, H., & Schmitz, C. (2002). Antioxidants and Alzheimer's disease: From bench to bedside (and back again). *Current Opinion in Clinical Nutrition and Metabolic Care, 5*, 645–651.

Scarmeas, N., Levy, G., Tang, M. X., Manly, J., & Stern, Y. (2001). Influence of leisure activity on the incidence of Alzheimer's disease. *Neurology, 57*, 2236–2242.

Scarmeas, N., & Stern, Y. (2003). Cognitive reserve and lifestyle. *Journal of Clinical and Experimental Neuropsychology, 25*, 625–633.

Schaie, K. W. (2005). *Developmental influences on adult intelligence: The Seattle Longitudinal Study*. New York, NY: Oxford University Press.

Schonfeld, L., King-Kallimanis, B. L., Duchene, D. M., Etheridge, R. L., Herrera, J. R., Barry, K. L., et al. (2010). Screening and brief intervention for substance misuse among older adults: The Florida BRITE project. *American Journal of Public Health, 100*, 108–114.

Schooler, C. (1987). Psychological effects of complex environments during the life-span: A review and theory. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 24–49). Norwood, NJ: Ablex.

Schooler, C., & Mulatu, M. S. (2001). The reciprocal effects of leisure time activities and intellectual functioning in older people: A longitudinal analysis. *Psychology and Aging, 16* (3), 466–482.

Schooler, C., Mulatu, M. S., & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging, 14*, 483–506.

Seeman, T. E., Lusignolo, T. M., Albert, M., & Berkman, L. (2001). Social relationships, social support, and patterns of cognitive aging in healthy, high-functioning older adults: MacArthur studies of successful aging. *Health Psychology, 20*, 243–255.

Shepherd, A. (2009). Nutrition through the life span, Part 3: Adults aged 65 years and over. *British Journal of Nursing, 18*, 301–307.

Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., et al. (2009). A cognitive training program based on principles of brain plasticity: Results from the Improvement in Memory with Plasticity-Based Adaptive Cognitive Training (IMPACT) study. *Journal of American Geriatrics Society, 57*, 594–603.

Smyer, M. A., & Qualls, S. H. (1999). *Aging and mental health*. Malden, MA: Blackwell.

Stafford, M., McMunn, A., Zaninottoa, P., & Nazroo, J. (2011). Positive and negative exchanges in social relationships as predictors of depression: Evidence from the English Longitudinal Study of Aging. *Journal of Aging Health, 23*, 607–628.

Stampfer, M. J., Kang, J. H., Chen, J., Cherry, R., & Grodstein, F. (2005). Effects of moderate alcohol consumption on cognitive function in women. *New England Journal of Medicine, 352*, 245–253.

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 09-443*. Rockville, MD: Author.

Syddall, H., Roberts, H. C., Evandrou, M., Cooper, C., Bergman, H., & Sayer, A. A. (2010). Prevalence and correlates of frailty among community-dwelling men and women: Findings from the Hertfordshire cohort study. *Age and Ageing, 39*, 197–203.

Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., et al. (2006). Neuropsychological prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. *Archives of General Psychiatry, 63*, 916–924.

Tan, E. J., Xue, Q.-L., Li, T., Carlson, M. C., & Fried, L. P. (2006). Volunteering: A physical activity intervention for older adults: The experience corps program in Baltimore. *Journal of Urban Health, 83*, 954–969.

Tang, M. X., Cross, P., Jacobs, D. M., Small, S., Bell, K., Merchant, C., et al. (2001). Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology, 56*, 49–56.

Topinková, E. (2008). Aging, disability and frailty. *Annuals of Nutrition and Metabolism, 52* (Suppl 1), 6–11.

Townsend, A., Noelker, L., Deimling, G., & Bass, D. (1989). Longitudinal impact of interhousehold caregiving on adult children's mental health. *Psychology and Aging, 4*, 393–401.

U.S. Census Bureau. (2011). *Statistical abstract of the United States: 2011* (130th ed.). Washington, DC: Author.

U.S. Department of Labor. (2007). *Aging baby boomers in a new workforce development system*. Washington, DC: Author.

VanItallie, T. B. (2003). Frailty in the elderly: Contributions of sarcopenia and visceral protein depletion. *Metabolism, 52*(10 Suppl 2), 22–26.

Verghese, J., Lipton, R. B., Katz, M. J., Hall, C. B., Derby, C. A., Kuslansky, G., et al. (2003). Leisure activities and the risk of dementia in the elderly. *New England Journal of Medicine, 348*, 2508–2516.

Wang, H. X., Karp, A., Winblad, B., & Fratiglioni, L. (2002). Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: A longitudinal study from the Kungsholmen project. *American Journal of Epidemiology, 155*, 1081–1087.

Waters, D. L., Baumgartner, R. N., Garry, P. J., & Vallas, B. (2010). Advantages of dietary, exercise-related, and therapeutic interventions to prevent and treat sarcopenia in adult patients: An update. *Clinical Interventions in Aging, 5*, 259–270.

Weber, J., & Wong, K. (2010). Older adults coping with vision loss. *Journal of Home Health Care Services Quarterly, 29*(3), 105–119.

Whitt-Glover, M. C., Taylor, W. C., Heath, G. W., & Macera, C. A. (2007). Self-reported physical activity among Blacks estimates from national surveys. *American Journal of Preventive Medicine, 33*, 412–417.

Willis, S. L., Tennstedt, S. L., Ball, K., Elias, J., Koepke, K. M., Morris, J. N., et al. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *Journal of American Medical Association, 296*, 2805–2814.

Wilson, R. S., & Bennett, D. A. (2003). Cognitive activity and risk of Alzheimer's disease. *Current Directions in Psychological Science, 12*, 87–91.

Xu, J. Q., Kockanek, K. D., Murphy, S. L., & Tejada-Vera, B. (2010). *Deaths: Final data for, 2007. National Vital Statistics Report, 58(19)*. Hyattsville, MD: National Center for Health Statistics.

Part II
Environmental Influences and Implications for Intervention Development

Chapter 6
Family

J. Mark Eddy, Charles R. Martinez Jr., Carol W. Metzler, and Richard E. Heyman

Introduction

Close relationships play vital roles in the development and maintenance of the physical and psychological health and illness of children and adults (Beach et al., 2006). Of particular importance are enduring relationships within families, including intimate relationships between adults and caregiving relationships among children, their parents, and other parental figures. Not only do positive family relationships promote resiliency across all stages of the life course, distressed family relationships are related to disability and suffering.

A particular research focus during the first several decades of the field of prevention science has been the impact of family relationships on children. A variety of social contexts have been identified, some internal and some external to families, that either buffer against or create risk for poor behavioral and mental health outcomes for children. Both quality of the parent–child relationship and quality of parenting practices have been shown to be associated with appraisals of family stress, adult adjustment (e.g., depression, quality of the couple relationship), and child adjustment (e.g., antisocial behavior, psychosocial well-being; Reid, Patterson, & Snyder, 2002). Although other contextual influences (e.g., socioeconomic status, neighborhood factors, family structure transitions, sociocultural factors, peer influences, school environment) have also been shown to influence outcomes, studies have shown that such factors often exert their primary effects through their impacts on parenting and the family environment (e.g., DeGarmo, Forgatch, & Martinez, 1999; Martinez & Forgatch, 2002; Stormshak, Connell, & Dishion, 2009).

While warm and nurturing interactions are important for positive adjustment across the lifespan (Biglan, Flay, Embry, & Sandler, 2012), specific aspects of

J.M. Eddy (✉)
Partners for Our Children, University of Washington, UW Tower, 18th Floor, Box 359476, Seattle, WA 98195-9476, USA
e-mail: jmarke@uw.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_6,

parent–child and other family relationships vary in their importance and influence during different periods of development. For example, during the first years of life, attentive, warm, and responsive parenting is central in the development of attachment, trust, and a healthy framework for the future relationships a child will form (Cassidy & Shaver, 2008). In contrast, during adolescence, close and consistent parental monitoring of the activities and the peer and adult relationships of a child is central to avoiding harm and fostering prosocial engagement (Racz & McMahon, 2011). Prevention scientists use basic knowledge such as this about child development and parenting to create tailored interventions intended to strengthen families with infants or families with teenagers. In this chapter, we introduce theoretical, correlational, and experimental bases that position the family as central to both the prevention of negative outcomes and the promotion of positive outcomes across the lifespan.

Theoretical Basis

Findings from a variety of studies indicate that a broad range of child problem behaviors are often initiated within a common developmental sequence (Reid & Eddy, 1997; Robertson, David, & Rao, 2003). Early child noncompliance at home has been shown to be linked to subsequent aggression and other conduct problems within the family (Chamberlain & Patterson, 1995; Patterson, Reid, & Dishion, 1992). In turn, conduct problems at home have been shown to predict behavioral and academic problems at school, engagement with other children who have similar problems (so-called deviant peers), and, later, high-risk health behaviors, such as substance use and sexual behavior in adolescence (Block, Block, & Keyes, 1988; Kellam, Brown, Rubin, & Ensminger, 1983). Thus, once begun, and without ongoing counters from adults and peers, one problem behavior often leads to another. This "co-occurrence" of youth problem behaviors has been documented both within general populations (Amaro, Blake, Schwartz, & Flinchbaugh, 2001; Biglan et al., 2004; Tapia, Schwartz, Prado, Lopez, & Pantin, 2006) and within racial and ethnic minority groups (Gil, Wagner, & Vega, 2000; Martinez & Eddy, 2005; National Institute on Drug Abuse, 1993; Turner & Hench, 2003) and has been hypothesized to be reflective of an underlying "general problem behavior" syndrome (Jessor & Jessor, 1977).

A number of theories have been proposed that bring together information about developmental sequence with information about the relations between and among risk and protective factors and child problem behaviors and competencies across childhood (e.g., Catalano, Hawkins, Berglund, Pollard, & Arthur, 2002; Coie & Jacobs, 1993; Kia-Keating, Dowdy, Morgan, & Noam, 2011; Schwartz, Pantin, Coatsworth, & Szapocznik, 2007). The focal points of these theories vary. Some emphasize the role that social bonds play (e.g., Hawkins & Weis, 1985) or the influence of social fields (e.g., Kellam & Rebok, 1992) as children grow into adults. Others emphasize the interplay between cognitions, affect, and behavior at specific

points in the lifespan (e.g., Bell & Naugle, 2008). Most theories, however, reflect key ideas from two major lineages, cognitive behavioral theory and ecological theory, and each highlights the importance of family relationships across the lifespan.

For example, social learning theory (e.g., Reid et al., 2002), a version of cognitive behavioral theory, posits that social contextual factors influence outcomes for children through their impact on the daily social interactional processes that occur within family relationships. Family members are presumed to influence one another's behavior in a bidirectional shaping process (e.g., parent to child, child to parent, parent to parent), but the theory emphasizes the importance of social interactions from parent to child when considering child outcomes. In short, parents and other adults are seen as the actors with responsibility and authority to make corrections when things are going wrong and as the key players to keep things going right.

Unfortunately, when negative contexts impinge on a family and impact parental functioning, parenting practices can suffer and child adjustment can be negatively affected. Thus, the effects of contextual factors on children's adjustment are hypothesized to be mediated directly by parent adjustment and parenting practices. Supporting this view, longitudinal studies grounded in social learning theory have demonstrated that social contextual stressors impact characteristics of the early family environment such as parent–child relationships, parental affection, parenting practices, and family conflict. In turn, these factors have been found to influence the development of child emotional and cognitive responses to environmental challenges (e.g., psychosocial stressors) that are predictive of emotional adjustment, both concurrently and later in life (Bruce, Fisher, Pears, & Levine, 2009; Cicchetti & Rogosch, 2007; DeCaro & Worthman, 2008a, 2008b; Dozier et al., 2006; Gunnar & Donzella, 2002; Pears & Fisher, 2005). The general tenants of social learning theory have been applied to a wide variety of interpersonal behaviors and relationships. For example, theories of the development and maintenance of intimate partner violence have expanded from an almost exclusive focus on concepts such as power and control to posit an array of both internal and external factors that may contribute to the occurrence of violence within a relationship (e.g., Bell & Naugle, 2008).

Ecodevelopmental theory (e.g., Pantin, Schwartz, Sullivan, Prado, & Szapocznik, 2004; Szapocznik & Coatsworth, 1999) broadens the framework of social learning by considering the relative contribution of interacting contextual "spheres" of influence within a child's social world (e.g., family, peers, school, neighborhood, community) and emphasizes the potential embedded nature of spheres (e.g., the world of a young child is embedded within family, family is embedded within neighborhood, neighborhood within community). The theory incorporates and expands the "ecological" framework of Bronfenbrenner (1979), who organized social contextual influences within interacting systems: (a) macrosystems, which reflect the broad social and philosophical ideals that define a particular culture (e.g., cultural and societal values); (b) exosystems, in which a child does not directly participate but that impact key social members of his or her

life (e.g., parents' social support); (c) mesosystems, which are the interactions between key members of the different contexts in which a child participates directly (e.g., parental monitoring of peers); and (d) microsystems, in which a child participates directly (e.g., family, school, and peers).

Unlike social learning theory, ecodevelopmental models are often entrenched in cultural context and explicitly identify cultural factors (e.g., acculturation processes, cultural identity, and cultural values and norms) that are linked to developmental pathways for children across systems (Kulis et al., 2005; Martinez, 2006; Martinez & Eddy, 2005; Pantin et al., 2004; Prado, Szapocznik, Maldonado-Molina, Schwartz, & Pantin, 2008; Szapocznik, Prado, Burlew, Williams, & Santisteban, 2007). The active consideration of cultural factors is still relatively new across the field of prevention at large (Castro, Barrera, & Martinez, 2004; Kumpfer, Alvarado, Smith, & Bellamy, 2002). However, the ongoing narrative around preventive theory and intervention now includes concepts such as youth "developmental assets" and "thriving" (rather than simply not engaging in problem behaviors; Hawkins, Catalano, & Miller, 1992; Lerner, 2001). Both cultural factors and family relationships are often highlighted as important not only in the formation of assets but also in the occurrence of "resiliency" in the face of risk exposure (Bradley, Whiteside, Mundfrom, & Casey, 1994; Kia-Keating et al., 2011; Schwartz et al., 2007) and ultimately to thriving throughout life.

Correlational Basis

Family interactions during childhood are related to long-term adjustment. Although studies of human biological mechanisms related to social relationships are limited, work with animal models has begun to reveal the role of the brain in the formation and maintenance of relationships (Lim & Young, 2006). Separate neurobiological mechanisms have been identified for social recognition, social motivation, social approach, and social bonding. Early experiences in relationships actually shape the brain, and these changes to the brain impact subsequent behavior, affect, and social interactions (Totah & Plotsky, 2006). Traumatic events such as maternal separation during critical developmental periods seem particularly important to later social functioning.

Similarly, in studies of humans, exposure to trauma, such as maltreatment during childhood, has been found to be related to psychopathology and other problems during adulthood (Felitti et al., 1998; Horwitz, Widom, Spatz McLaughlin, & White, 2001; McCauley, Kern, Kolodner, Dill, & Schroeder, 1997). Maltreatment per se is most likely to occur very early in life, from birth to age 3 (U.S. Department of Health and Human Services, 2005). Experiencing maltreatment is also related to difficulties in relationships during adulthood, including both perpetration of and victimization from intimate partner violence (Capaldi, Knoble, Shortt, & Kim, 2012). However, this effect appears to be mediated by the relationship between childhood abuse and the development of ongoing child problem behaviors, most

notably aggression, noncompliance, and other antisocial behaviors that occur across settings (e.g., home, school, community), as well as substance use and abuse (White & Widom, 2003), which then increase an individual's risk for involvement in intimate partner violence. Similar relationships have been found between witnessing intimate partner violence during childhood (Hungerford, Wait, Fritz, & Clements, 2012), the development of antisocial behavior, and perpetration of and victimization by intimate partner violence during adulthood (Moffitt, Krueger, Caspi, & Fagan, 2000).

Adult intimate relationships also play a key role in the physical and emotional well-being of the involved partners (Glen & Weaver, 1981). Over and above age and income levels, both men and women in marital relationships tend to be healthier than separated, divorced, or single individuals (Johnson, Backlund, & Sorlie, 2000). Similar findings are emerging for men and women in cohabiting relationships (Wu, Penning, Pollard, & Hart, 2003). Adjusting for demographic factors, the risk of mortality for married or cohabiting adults is also lower (Gordon & Rosenthal, 1995). Although the mechanisms for these effects have not yet been clearly identified, married individuals, for example, do tend to have better immune function than separated or divorced individuals (Kiecolt-Glaser, Fisher, Ogrocki, & Stout, 1987; Kiecolt-Glaser, Kennedy, Malkoff, & Fisher, 1988; see also Uchino, Cacioppo, & Kiecolt-Glaser, 1996). Unfortunately, marriage and similar enduring adult intimate relationships also can have a variety of serious negative effects beyond intimate partner violence. Being in an unhappy, discordant marital or cohabiting relationship is related to a variety of negative outcomes for individuals and families at large, such as poor health (Kiecolt-Glaser & Newton, 2001), depression (Weissman, 1987), other psychiatric disorders (Goering, Lin, Campbell, & Boyle, 1996), work-related disability (Appelberg, Romanov, Heikkila, Honkasalo, & Koskenvuo, 1996), and negative impacts on children (Cummings & Davies, 2010).

Ongoing discord in other family relationships, most notably parent–child relationships, also plays a role in the development of child problem behaviors, particularly antisocial behaviors (Loeber & Dishion, 1983; Patterson, 1982). Although there may be a variety of causes related to the genesis of antisocial behaviors, key factors in their maintenance and escalation are inept parenting practices (Patterson et al., 1992). Parenting practices that appear to be particularly important in this regard include inconsistent discipline, inadvertent reinforcement of aggression and other aversive behavior, and lack of monitoring and supervision (Capaldi, Chamberlain, & Patterson, 1997; Racz & McMahon, 2011).

The development of child antisocial behaviors and, conversely, the development of prosocial behaviors are particularly important outcomes in terms of health promotion and prevention. This is not only because of the intervening roles antisocial behaviors appear to play between childhood traumatic experiences and adult problems but also because of the relationship among antisocial behaviors, other significant problems during adolescence (such as substance use and high-risk sexual behavior), and poor outcomes during adulthood (Lipsey & Derzon, 1998; Reid et al., 2002). Specific types of parenting behaviors found to buffer against

problem development and foster prosocial functioning include the regular employment of parental encouragement of positive child behaviors, the use of consistent and nonharsh consequences when inappropriate behaviors occur, the practice of close monitoring and supervision so that the occurrence of both positive and negative behaviors are noticed, and the engagement by caregivers in advocacy for meeting the specific needs of a child, such as within the school environment (Kumpfer & Alder, 2003).

Experimental Basis

Findings such as those discussed above have been influential not only in the development of theories in prevention science but also in the development of preventive interventions. Over the past few decades, multiple experimental tests of theoretically informed family-based preventive interventions have been conducted. The most scientifically rigorous of these have been randomized trials, in which a sample is drawn from an identified population, interested participants are randomly assigned to receive an intervention of interest or to a comparison condition (e.g., a no-intervention "control" group), the intended intervention is delivered as planned, and child outcomes are tracked across time. When such trials meet high-quality standards, they can be valid examinations of whether an intervention caused the observed outcomes and, in the case of a family-based intervention trial, whether changing something in a family results in changes in child outcomes. The findings from these randomized trials lend support for the importance of families as important agents in preventive efforts, particularly in the area of child problem behaviors. Similar studies on health promotion are fewer in number.

Not surprisingly, given the role such behavior plays in the genesis of other problems, one of the most frequently studied childhood problems is antisocial behavior. Through what is available via electronic databases alone, Grove, Evans, Pastor, and Mack (2008) identified almost 65,000 studies of child antisocial behavior. It is clear from this body of work that individuals who display high levels of antisocial behaviors during childhood ultimately invoke significant costs not only to their victims but also to their families, to themselves, and to society. Scott, Knapp, Henderson, and Maughan (2001), for example, estimated that in adulthood, men who exhibit persistent, frequent, and serious antisocial behaviors during their teenage years generate costs to society that are more than ten times those of men who did not exhibit such problems. These costs often begin with school counselor involvement and special education programs in school; progress to contact with police, juvenile justice, residential care, and mental health and substance abuse treatment programs; and continue to chronic involvement with the various facets of the adult criminal justice system, including police, jails, prisons, and probation and parole departments.

Within the numerous studies they identified, Grove and colleagues (2008) found 45 recent, high-quality randomized trials of prevention programs that focused on

child antisocial behavior and that included at least 6 months of follow-up following the end of the program of interest. These encompassed a total of 9,366 undiagnosed child participants and yielded 91 tests of differences in outcomes between intervention and comparison groups. A common index of the magnitude of difference is called "effect size," with positive values indicating that the intervention of interest is related to relatively better outcomes for participants, 0 indicating no difference, and negative values indicating worse outcomes. The greater the magnitude of the effect size, the stronger the difference between the groups.

In their meta-analysis of the findings across these trials, Grove and colleagues (2008) found that the average effect size 2 years after the end of intervention was +0.17, indicating a small impact of preventive interventions on child antisocial behavior. Behavioral programs appeared to have the strongest impact (+0.54), followed by multiple-component programs (+0.21), cognitive behavioral programs (+0.13), eclectic programs (+0.03), and cognitive programs (−0.09). Family-based programs were most commonly classified as "behavioral" or "cognitive behavioral," and most multiple-component programs included a family-based component. Thus, there is experimental support, across repeated studies, for family-based preventive interventions inhibiting the development of child antisocial behavior.

Most such interventions are either some version of or heavily influenced by a particular class of interventions called parent management training (PMT). PMT teaches and then coaches parents to use cognitive behavioral techniques to encourage child prosocial behaviors and discourage child antisocial behaviors (Patterson, Reid, Jones, & Conger, 1975). Programs typically focus on the development of a broad repertoire, or "toolbox," of parenting skills in the areas of positive reinforcement; nonviolent, nonharsh, and consistent discipline; effective monitoring and supervision; and constructive family problem-solving. In an early meta-analysis of randomized trials that compared outcomes of a PMT versus a comparison group with youth already exhibiting clinical levels of antisocial behavior, the average effect size for antisocial outcomes immediately post-intervention was +0.80 (Serketich & Dumas, 1996; see also Brestan & Eyberg, 1998), indicating a much stronger impact than preventive efforts. More broadly, a recent summary of findings from 26 meta-analyses of clinical treatments for children with antisocial behavior problems found an average effect size of +0.49 for behavioral/cognitive behavioral therapies (including PMT) and +0.41 for family therapies (Litschge, Vaughn, & McCrea, 2010).

Collectively, these findings suggest the value of family-based interventions in addressing child antisocial behavior and indicate that preventive effects might be stronger with higher-risk subgroups of children, an issue that Grove and colleagues (2008) and most meta-analyses have not investigated. For example, Grove and colleagues reported an effect size of +0.37 for the multiple-component program Linking the Interests of Families and Teachers (LIFT; Eddy, Reid, Stoolmiller, & Fetrow, 2003). LIFT included a PMT component, a child social skills and problem-solving training component (Losel & Beelmann, 2003), and a playground Good Behavior Game component (Dolan et al., 1993). A key finding for LIFT not examined in the Grove analysis was that some of the most significant impacts

occurred for the highest risk boys and girls in the sample (Reid, Eddy, Fetrow, & Stoolmiller, 1999). Most notable was the impact of the program on aggressive behavior on the playground. Children with very high levels of aggression before the program began were impacted much more strongly than children with high, average, and low levels of aggression before the beginning of the program (i.e., effect sizes of +0.79, +0.51, +0.23, and −0.05, respectively; Stoolmiller, Eddy, & Reid, 2000). When adjusted for low reliability and censoring in the measures of aggression, these effect sizes were estimated to be many times greater in magnitude, respectively, +5.13, +2.85, +0.57, and −1.71.

In the broader prevention literature, meta-analyses of effects related to preventive interventions targeting a wide variety of problems in childhood and adolescence beyond, but not excluding, child antisocial behaviors, have also found positive impacts of family-based programs. For example, in their comprehensive review of studies of "primary" prevention programs targeting any type of child outcome, Durlak and Wells (1997) found that programs targeting first-time mothers had one of the highest effect sizes (+0.87; nurse home visitation, Olds, Henderson, Chambelin, & Tatelbaum, 1986). The effect size for early childhood programs, which typically include a parent support component, was +0.35, and the effect size for parent training programs in general was +0.16. Exemplars of a range of family-based preventive programs for universal and at-risk populations, and their effects on child outcomes, are described in Metzler, Eddy, and Lichtenstein (2013).

Closing Comments

Results from correlational and experimental studies suggest that family relationships play an influential role in child development. Accordingly, family relationships and the contexts surrounding them are important drivers of children's outcomes within the theories that have guided and continue to guide the development of most current preventive interventions for children. Antisocial behaviors have received particular attention from researchers and have been found to be related to a wide variety of other serious problems, including academic failure, substance use, and early sexual behavior. The available evidence suggests that one effective way to support the positive development of children and to inhibit the growth of antisocial and related problem behaviors is to promote physically and emotionally nurturing family environments, including strong relationships between caregivers and children and the use of effective parenting practices (Biglan et al., 2012).

Although much has been learned from a scientific point of view about families, there are notable gaps in the knowledge base. For example, even though many studies have been conducted on the prevention of child problems, such as antisocial behavior, very little information has been gathered on the mechanisms through which successful prevention programs effect change. It appears that family-based prevention programs can reliably impact subsequent child problem behaviors.

However, except for a few studies (e.g., DeGarmo, Eddy, Reid, & Fetrow, 2009), it is unclear whether programs do this by changing parenting behaviors, by modifying other relevant family factors, and/or by doing something else altogether. In addition, although child problem behaviors, mental health, and physical health are linked, except for a few studies (e.g., Fisher, Gunnar, Dozier, Bruce, & Pears, 2006), it is unknown whether and how prevention programs that are intended to decrease child problem behaviors impact child physical health.

Even though more expansive ecologically based theories have become quite popular within the field of prevention, many aspects of these theories remain untested, and notable key components have yet to be integrated into most preventive interventions. For example, although these theories often hypothesize the importance of key aspects of human social structure such as culture, most basic and applied prevention research with families has been conducted within relatively limited cultural contexts, namely, Western, high-income countries. Further, most family-based interventions, even with cultural minority groups within those countries, do not take into account cultural factors.

There is a growing evidence base regarding the positive impact of culture on intervention that suggests that the omission of culture is to the peril of prevention efforts. To date, findings indicate that programs that take culture into account result in improvements in intervention participation and completion rates, as well as in the sustainability of interventions within communities (Catalano et al., 1993; Smith, Rodríguez, & Bernal, 2011). In a meta-analysis of outcomes from 76 studies of mental health-related interventions, Griner and Smith (2006) found that interventions that took culture into account performed quite favorably compared with standard interventions. A body of such work has yet to be developed across the prevention field.

Thus, although much has been learned about families, the contexts within which they function, and promising prevention strategies, there is no shortage of questions to answer. Studies of all types (i.e., qualitative, correlational, experimental) are needed in low- and middle-income countries around the world to better understand child development and family relationships within a wide variety of cultures. Studies are needed across the board on the mechanisms of change in family-based preventive interventions. Studies are needed that examine how family-based preventive interventions impact not only behaviors and cognitions but also physical health. Studies are needed of the impact of family-based interventions not only on reducing problem behaviors but on promoting prosocial behaviors that lead to success during childhood and adulthood. As the results from studies such as these accumulate, the theoretical models that guide prevention and health promotion should improve in their validity and specificity, and the interventions grounded in these theories can better assist families and children in thriving within the cultural milieu that they call “home.”

References

Amaro, H., Blake, S. M., Schwartz, P. M., & Flinchbaugh, L. J. (2001). Developing theory-based substance abuse prevention programs for young adolescent girls. *Journal of Early Adolescence, 21*, 256–293.

Appelberg, K., Romanov, K., Heikkila, K., Honkasalo, K., & Koskenvuo, M. (1996). Interpersonal conflict as a predictor of work disability: A follow-up study of 15,348 Finnish employees. *Journal of Psychosomatic Research, 40*, 157–167.

Beach, S. R. H., Wamboldt, M. Z., Kaslow, N. J., Heyman, R. E., First, M. B., Underwood, L. G., et al. (Eds.). (2006). *Relational processes and DSM-V: Neuroscience, assessment, prevention, and treatment*. Arlington, VA: American Psychiatric Publishing.

Bell, K. M., & Naugle, A. E. (2008). Intimate partner violence theoretical considerations: Moving towards a contextual framework. *Clinical Psychology Review, 28*, 1096–1107.

Biglan, A., Brennan, P. A., Foster, S. L., Holder, H. D., Miller, T. L., Cunningham, P. B., et al. (2004). *Helping adolescents at risk: Prevention of multiple problem behaviors*. New York, NY: Guilford.

Biglan, A., Flay, B. R., Embry, D. D., & Sandler, I. N. (2012). The critical role of nurturing environments for promoting human well-being. *American Psychologist, 67*(4), 257–271.

Block, J., Block, J. H., & Keyes, S. (1988). Longitudinally foretelling drug use in adolescence: Early childhood personality and environmental precursors. *Child Development, 59*, 336–355.

Bradley, R. H., Whiteside, L., Mundfrom, D. J., & Casey, P. H. (1994). Early indications of resilience and their relation to experiences in the home environments of low birthweight, premature children living in poverty. *Child Development, 65*, 346–360.

Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology, 27*, 180–189.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Bruce, J., Fisher, P. A., Pears, K. C., & Levine, S. (2009). Morning cortisol levels in preschool-aged foster children: Differential effects of maltreatment type. *Developmental Psychobiology, 51*(1), 14–23.

Capaldi, D. M., Chamberlain, P., & Patterson, G. R. (1997). Ineffective discipline and conduct problems in males: Association, late adolescent outcomes, and prevention. *Aggression and Violent Behavior, 2*, 343–353.

Capaldi, D. M., Knoble, N. B., Shortt, J. W., & Kim, H. K. (2012). A systematic review of risk factors for intimate partner violence. *Partner Abuse: New Directions in Research, Intervention, and Policy, 3*(2), 231–280.

Cassidy, J., & Shaver, P. R. (2008). *Handbook of attachment: Theory, research, and clinical applications* (2nd ed.). New York, NY: Guilford Press.

Castro, F. G., Barrera, M., Jr., & Martinez, C. R., Jr. (2004). The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prevention Science, 5*(1), 41–45.

Catalano, R. F., Hawkins, J. D., Berglund, L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*(Suppl. 6), 230–239.

Catalano, R. F., Hawkins, J. D., Krenz, C., Gillmore, M., Morrison, D., Wells, E., et al. (1993). Using research to guide culturally appropriate drug abuse prevention. *Journal of Consulting and Clinical Psychology, 61*(5), 804–811.

Chamberlain, P., & Patterson, G. R. (1995). Discipline and child compliance in parenting. In M. H. Bornstein (Ed.), *Handbook of parenting* (Vol. 4, pp. 202–225). Mahwah, NJ: Lawrence Erlbaum Associates.

Cicchetti, D., & Rogosch, F. A. (2007). Personality, adrenal steroid hormones, and resilience in maltreated children: A multilevel perspective. *Development and Psychopathology, 19*, 787–809.

Coie, J. D., & Jacobs, M. R. (1993). The role of social context in the prevention of conduct disorder. *Development and Psychopathology, 5*, 263–275.

Cummings, E. M., & Davies, P. T. (2010). *Marital conflict and children: An emotional security perspective*. New York, NY: The Guilford Press.

DeCaro, J. A., & Worthman, C. M. (2008a). Culture and the socialization of child cardiovascular regulation at school entry in the U.S. *American Journal of Human Biology, 20*, 572–583.

DeCaro, J. A., & Worthman, C. M. (2008b). Return to school accompanied by changing associations between family ecology and cortisol. *Developmental Psychobiology, 50*, 183–195.

DeGarmo, D. S., Eddy, J. M., Reid, J. B., & Fetrow, R. A. (2009). Evaluating mediators of the impact of the linking the interests of families and teachers (LIFT) multimodal preventive intervention on substance use initiation and growth across adolescence. *Prevention Science, 10*, 208–220.

DeGarmo, D. S., Forgatch, M. S., & Martinez, C. R., Jr. (1999). Parenting of divorced mothers as a link between social status and boys' academic outcomes: Unpacking the effects of SES. *Child Development, 70*(5), 1231–1245.

Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., et al. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology, 14*, 317–345.

Dozier, M., Manni, M., Gordon, M. K., Peloso, E., Gunnar, M. R., Stovall-McClough, K. C., et al. (2006). Foster children's diurnal production of cortisol: An exploratory study. *Child Maltreatment, 11*, 189–197.

Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology, 25*(2), 115–152.

Eddy, J. M., Reid, J. B., Stoolmiller, M., & Fetrow, R. A. (2003). Outcomes during middle school for an elementary school-based preventive intervention for conduct problems: Follow-up results from a randomized trial. *Behavior Therapy, 34*, 535–552.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., et al. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) study. *American Journal of Preventive Medicine, 14*(4), 245–258.

Fisher, P. A., Gunnar, M., Dozier, M., Bruce, J., & Pears, K. C. (2006). Effects of a therapeutic intervention for foster children on behavior problems, caregiver attachment, and stress regulatory neural systems. *Annals of the New York Academy of Sciences, 1094*, 215–225.

Gil, A. G., Wagner, E. F., & Vega, W. A. (2000). Acculturation, familism, and alcohol use among Latino adolescent males: Longitudinal relations. *Journal of Community Psychology, 28*(4), 443–458.

Glen, N. D., & Weaver, C. N. (1981). The contribution of martial happiness to global happiness. *Journal of Marriage and the Family, 43*, 161–168.

Goering, P., Lin, E., Campbell, D., & Boyle, M. H. (1996). Psychiatric disability in Ontario. *Canadian Journal of Psychiatry, 41*, 564–571.

Gordon, H. S., & Rosenthal, G. E. (1995). Impact of marital status on outcomes in hospitalized patients. *Archives of Internal Medicine, 155*, 2465–2471.

Griner, D., & Smith, T. B. (2006). Culturally adapted mental health intervention: A meta-analytic review. *Psychotherapy: Theory, Research, Practice, Training, 43*(4), 531–548.

Grove, A. B., Evans, S. W., Pastor, D. A., & Mack, S. D. (2008). A meta-analytic examination of follow-up studies of programs designed to prevent the primary symptoms of oppositional defiant and conduct disorders. *Aggression and Violent Behavior, 13*, 169–184.

Gunnar, M., & Donzella, B. (2002). Social regulation of the cortisol levels in early human development. *Psychoneuroendocrinology, 27*, 199–220.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112*, 64–105.

Hawkins, J. D., & Weis, J. G. (1985). The social development model: An integrated approach to delinquency prevention. *Journal of Primary Prevention, 6*(2), 73–95.

Horwitz, A. V., Widom, C. S., Spatz McLaughlin, J., & White, H. R. (2001). The impact of childhood abuse and neglect on adult mental health: A prospective study. *Journal of Health and Social Behavior, 42*, 184–201.

Hungerford, A., Wait, S. K., Fritz, A. M., & Clements, C. M. (2012). Exposure to intimate partner violence and children's psychological adjustment, cognitive functioning, and social competence: A review. *Aggression and Violent Behavior, 17*, 373–382.

Jessor, R., & Jessor, S. L. (1977). *Problem behavior and psychosocial development*. New York, NY: Academic.

Johnson, N. J., Backlund, E., & Sorlie, P. D. (2000). Marital status and mortality: The National Longitudinal Mortality Study. *Annals of Epidemiology, 10*, 224–238.

Kellam, S. G., Brown, C. H., Rubin, B. R., & Ensminger, M. E. (1983). Paths leading to teenage psychiatric symptoms and substance use: Developmental epidemiological studies in Woodlawn. In S. B. Guze, F. J. Earls, & J. E. Barrett (Eds.), *Childhood psychopathology and development* (pp. 17–47). New York, NY: Raven Press.

Kellam, S. G., & Rebok, G. W. (1992). Building developmental and etiological theory through epidemiologically based preventive intervention trials. In J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 162–195). New York, NY: Guilford Press.

Kia-Keating, M., Dowdy, E., Morgan, M. L., & Noam, G. G. (2011). Protecting and promoting: An integrative conceptual model for healthy development of adolescents. *Journal of Adolescent Health, 48*, 220–228.

Kiecolt-Glaser, J. K., Fisher, L. D., Ogrocki, P., & Stout, J. C. (1987). Marital quality, marital disruption, and immune function. *Psychosomatic Medicine, 49*, 31–34.

Kiecolt-Glaser, J. K., Kennedy, S., Malkoff, S., & Fisher, L. (1988). Marital discord and immunity in males. *Psychosomatic Medicine, 50*, 213–299.

Kiecolt-Glaser, J. K., & Newton, T. (2001). Marriage and health: His and hers. *Psychological Bulletin, 127*, 472–503.

Kulis, S., Marsiglia, F. F., Elek, E., Dustman, P., Wagstaff, D. A., & Hecht, M. (2005). Mexican/Mexican American adolescents and Keepin' it REAL: An evidence-based substance use prevention program. *Children and Schools, 27*(3), 133–145.

Kumpfer, K. L., & Alder, S. (2003). Dissemination of research-based family interventions for the prevention of substance abuse. In Z. Sloboda & W. J. Bukoski (Eds.), *Handbook of drug abuse prevention* (pp. 75–100). New York, NY: Kluwer Academic/Plenum.

Kumpfer, K. L., Alvarado, R., Smith, P., & Bellamy, N. (2002). Cultural sensitivity and adaptation in family-based prevention interventions. *Prevention Science, 3*, 241–246.

Lerner, R. M. (2001). Promoting promotion in the development of prevention science. *Applied Developmental Science, 5*(4), 254–257.

Lim, M. M., & Young, L. J. (2006). Neurobiology of the social brain: Lessons from animal models about social relationships. In S. R. H. Beach, M. Z. Wamboldt, N. J. Kaslow, R. E. Heyman, M. B. First, L. G. Underwood, & D. Reiss (Eds.), *Relational processes and DSM-V: Neuroscience, assessment, prevention, and treatment* (pp. 21–38). Arlington, VA: American Psychiatric Publishing.

Lipsey, M. W., & Derzon, J. H. (1998). Predictors of violent or serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 86–105). Thousand Oaks, CA: Sage.

Litschge, C. M., Vaughn, M. G., & McCrea, C. (2010). The empirical status of treatments for children and youth with conduct problems: An overview of meta-analytic studies. *Research on Social Work Practice, 20*(1), 21–35.

Loeber, R., & Dishion, T. J. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin, 94*, 68–99.

Losel, F., & Beelmann, A. (2003). Effects of child skills training in preventing antisocial behavior: A systematic review of randomized evaluations. *Annals of the American Academy of Political and Social Science, 587*, 84–109.

Martinez, C. R., Jr. (2006). Effects of differential family acculturation on Latino adolescent substance use. *Family Relations, 55*(3), 306–317.

Martinez, C. R., Jr., & Eddy, J. M. (2005). Effects of culturally adapted parent management training on Latino youth behavioral health outcomes. *Journal of Consulting and Clinical Psychology, 73*(4), 841–851.

Martinez, C. R., Jr., & Forgatch, M. S. (2002). Adjusting to change: Linking family structure transitions with parenting and boys' adjustment. *Journal of Family Psychology, 16*(2), 107–117.

McCauley, J., Kern, D. E., Kolodner, K., Dill, L., & Schroeder, A. F. (1997). Clinical characteristics of women with a history of childhood abuse: Unhealed wounds. *Journal of the American Medical Association, 277*, 1362–1368.

Metzler, C. W., Eddy, M., & Lichtenstein, D. (2013). Prevention approaches aimed at individuals. In B. McCrady & E. Epstein (Eds.), *Addictions: A comprehensive guidebook* (2nd ed.). New York, NY: Oxford University Press.

Moffitt, T. E., Krueger, R. F., Caspi, A., & Fagan, J. (2000). Partner abuse and general crime: How are they the same? How are they different? *Criminology, 38*(1), 199–232.

National Institute on Drug Abuse. (1993). *Drug use among racial/ethnic minorities (NIH Pub No. 03-3888)*. Bethesda, MD: Division of Epidemiology, Services & Prevention Research, National Institutes of Health, U.S. Department of Health and Human Services.

Olds, D. L., Henderson, C. R., Chambelin, R., & Tatelbaum, R. (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics, 78*, 65–78.

Pantin, H., Schwartz, S. J., Sullivan, S., Prado, G., & Szapocznik, J. (2004). Ecodevelopmental HIV prevention programs for Hispanic adolescents. *American Journal of Orthopsychiatry, 74* (4), 545–558.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *A social learning approach: IV, Antisocial boys*. Eugene, OR: Castalia.

Patterson, G. R., Reid, J. B., Jones, R. R., & Conger, R. E. (1975). *A social learning approach to family intervention: Families with aggressive children* (Vol. 1). Eugene, OR: Castalia Publishing.

Pears, K., & Fisher, P. A. (2005). Developmental, cognitive, and neuropsychological functioning in preschool-aged foster children: Associations with prior maltreatment and placement history. *Journal of Developmental and Behavioral Pediatrics, 26*, 112–122.

Prado, G., Szapocznik, J., Maldonado-Molina, M. M., Schwartz, S. J., & Pantin, H. (2008). Drug use/abuse prevalence, etiology, prevention, and treatment in Hispanic adolescents: A cultural perspective. *Journal of Drug Issues, 28*(1), 5–36.

Racz, S. J., & McMahon, R. J. (2011). The relationship between parental knowledge and monitoring and child and adolescent conduct problems: A 10-year update. *Clinical Child and Family Psychology Review, 14*(4), 377–398.

Reid, J. B., & Eddy, J. M. (1997). The prevention of antisocial behavior: Some considerations in the search for effective interventions. In D. M. Stoff, J. Breiling, & J. D. Master (Eds.), *The handbook of antisocial behavior* (pp. 343–356). New York, NY: Wiley.

Reid, J. B., Eddy, J. M., Fetrow, R. A., & Stoolmiller, M. (1999). Description and immediate impacts of a preventative intervention for conduct problems. *American Journal of Community Psychology, 27*, 483–517.

Reid, J. B., Patterson, G. R., & Snyder, J. (Eds.). (2002). *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention*. Washington, DC: American Psychological Association.

Robertson, E. B., David, S. L., & Rao, S. A. (2003). *Preventing drug abuse among children and adolescents (NIH Publication No. 04-4212)* (2nd ed.). Bethesda, MD: National Institute on Drug Abuse.

Schwartz, S. J., Pantin, H., Coatsworth, J. D., & Szapocznik, J. (2007). Addressing the challenges and opportunities for today's youth: Toward an integrative model and its implications for research and intervention. *Journal of Primary Prevention, 28*(2), 117–144.

Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: Follow up study of antisocial children into adulthood. *British Medical Journal, 323*, 191.

Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27*, 171–186.

Smith, T. B., Rodríguez, M. D., & Bernal, G. (2011). Culture. *Journal of Clinical Psychology, 67* (2), 166–175.

Stoolmiller, M., Eddy, J. M., & Reid, J. B. (2000). Detecting and describing preventative intervention effects in a universal school-based randomized trail targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology, 68*, 296–306.

Stormshak, E. A., Connell, A., & Dishion, T. J. (2009). An adaptive approach to family-centered intervention in schools: Linking intervention engagement to academic outcomes in middle and high school. *Prevention Science, 10*, 221–235.

Szapocznik, J., & Coatsworth, J. D. (1999). An ecodevelopmental framework for organizing the influences on drug abuse: A developmental model of risk and protection. In M. D. Glantz & C. R. Hartel (Eds.), *Drug abuse: Origins & interventions* (pp. 331–366). Washington, DC: American Psychological Association.

Szapocznik, J., Prado, G., Burlew, A. K., Williams, R. A., & Santisteban, D. A. (2007). Drug abuse in African American and Hispanic adolescents: Culture, development and behavior. *Annual Review of Clinical Behavior, 3*, 77–105.

Tapia, M. I., Schwartz, S. J., Prado, G., Lopez, B., & Pantin, H. (2006). Parent-centered intervention: A practical approach for preventing drug abuse in Hispanic adolescents. *Research on Social Work Practice, 16*, 146–165.

Totah, N. K. B., & Plotsky, P. M. (2006). Refining the categorical landscape of the DSM: Role of animal models. In S. R. H. Beach, M. Z. Wamboldt, N. J. Kaslow, R. E. Heyman, M. B. First, L. G. Underwood, & D. Reiss (Eds.), *Relational processes and DSM-V: Neuroscience, assessment, prevention, and treatment* (pp. 39–60). Arlington, VA: American Psychiatric Publishing.

Turner, W. L., & Hench, M. (2003). Drug abuse prevention among African Americans. In Z. Sloboda & W. Bukoski (Eds.), *Handbook for drug abuse prevention, theory, science, and practice*. Baltimore, MD: Williams & Wilkins.

U.S. Department of Health and Human Services. (2005). *Children's bureau: Child maltreatment, 2003*. Washington, DC: U.S. Government Printing Office.

Uchino, B. N., Cacioppo, J. T., & Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin, 119*, 488–531.

Weissman, M. M. (1987). Advances in psychiatric epidemiology: Rates and risks for major depression. *American Journal of Public Health, 77*, 445–551.

White, H. R., & Widom, C. S. (2003). Intimate partner violence among abused and neglected children in young adulthood: The mediating effects of early aggression, antisocial personality, hostility and alcohol problems. *Aggressive Behavior, 29*, 332–345.

Wu, Z., Penning, M. J., Pollard, M. S., & Hart, R. (2003). In sickness and in health: Does cohabitation count. *Journal of Family Issues, 24*, 811–838.

Chapter 7
School Influences on Child and Youth Development

David Osher, Kimberly Kendziora, Elizabeth Spier, and Mark L. Garibaldi

A Historical View of the Function and Organization of the School as a Social Institution

Schools have played a key role in youth development throughout American history, both serving as social microcosms of the broader society and reciprocally influencing communities and their people (Kidron & Osher, 2012; Rury, 2002). Educational events have often become cultural milestones, such as the contested integration of schools following *Brown v. the Board of Education of Topeka* in 1954, the establishment of Head Start in 1965, and the mandate to provide a free and appropriate education to all students through the Education for All Handicapped Children Act of 1975. In more recent years, the influences of public accountability for the achievement of all subgroups of students under No Child Left Behind, the expansion of charter schools and school choice initiatives, and the special focus on persistently low-performing schools under Race to the Top have influenced not only how communities engage with their school systems but also the kinds of educational outcomes students attain.

The importance of schools to youth development was emphasized by John Dewey and is well specified in life course/social field theory (Kellam, Branch, Agrawal, & Ensminger, 1975). In this perspective, one or more main social fields are critically important in each stage of life, and in each social field there are defined social task demands. Success or failure in regard to social roles is marked by the adequacy of behavioral responses of each individual to the specific social task demands faced within each main social field at each stage of life. The social task demands are defined by, and the adequacy of responses are rated by, *natural raters*, such as parents in the home, teachers in the classroom, or significant peers in the

D. Osher (✉)
Health and Social Development Program, American Institutes for Research (AIR), 1000 Thomas Jefferson St, NW, Washington, DC 20007, USA
e-mail: dosher@air.org

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_7,

peer group. Not only the individual's performance but also chance and the fit of the individual in the social context play roles in success or failure. Aggression, academic problems, and other early antecedents to problem outcomes must be viewed, therefore, as residing not merely in the child, but also as the social fields of family, school, and community.

Within the social field of the school, students confront major developmental challenges, negotiate group and intergroup social relationships, and acquire (or fail to acquire) important capacities that may enable them to thrive (Cairns & Cairns, 1994). From the perspective of risk and protection, schools may function as a protective factor, creating a safe harbor, offering both challenge and a sense of mission, fostering positive relationships with adults and peers, developing competencies and a sense of efficacy, and providing students with access to social capital, mental health supports, and leadership opportunities. Unfortunately, schools may also be stressful places that function as a risk factor, as youth adapt their behaviors to relatively inflexible bureaucratic structures and adult-driven demands within a high-stakes environment (e.g., Eccles & Midgley, 1989). Instead of safe harbors, schools can expose students to physical and emotional violence, boredom, alienation, academic frustration, negative relationships with adults and peers, teasing, bullying, gangs, humiliation and failure, harsh punishment, and expulsion from the school community and its resources.

In this chapter, we examine schools as a social institution, examining relevant student-, teacher-, and building-level characteristics that are particularly salient to prevention scientists. We also briefly describe schools as a locus for preventive interventions.

Student-Level Predictors of Academic and Social Development

A plethora of developmental research addresses student-level competencies and academic achievement. In this section, we review those student-level variables that interact with individual students' social, emotional, cognitive, and behavioral competencies: (a) cognitive skills, (b) self-regulation, and (c) social and emotional competencies.

Cognitive/Learning Skills

The central importance of cognitive skills for academic achievement cannot be ignored. Using well-established IQ tests that are culturally and linguistically appropriate, the magnitude of reported concurrent IQ-achievement correlations is consistently in the range of 0.60–0.80 (Sattler & Dumont, 2004; Watkins, Lei, &

Canivez, 2007). The strength of this association is among the strongest reported across all fields of psychology.

Intelligence, however, is understood to be a multifaceted trait that can be strongly influenced by the environment. The specific cognitive skills involved in learning (although there are many diverse models and lists) generally consist of attention, perception, memory, and reasoning. As students get older, problem-solving, higher-order and critical thinking, analysis, and synthesis skills also become important for postsecondary success. Additionally, *metacognition* refers to awareness of and the exertion of active control over the specific cognitive functions associated with learning (Kuhn, 2000; Nelson & Narens, 1994). Teaching practices congruent with a metacognitive approach to learning include those that focus on sense-making, self-assessment, and reflection on what worked and what needs improving. Sense-making in particular can be subjective and culture bound, but understanding the way students make sense of things can help them break through achievement barriers. Metacognitive teaching practices have been shown to increase the degree to which students transfer their learning to new settings and events (Bransford, Brown, & Cocking, 1999; Wirkala & Kuhn, 2011).

Children who are raised in economically disadvantaged neighborhoods start kindergarten with significantly lower cognitive skills than middle-class children do (Lee & Burkam, 2002; Shonkoff & Phillips, 2000). Before starting school, the average cognitive score of children in the highest socioeconomic group is 60 % higher than that of children in the lowest group. Socioeconomic status is by far the most powerful factor influencing the results of tests of cognitive skills. Factors that affect this include family educational expectations, exposure to language and reading, child care quality, computer use, and television habits (e.g., Hart & Risley, 1995). High-quality early childhood education for poor children does produce long-lasting improvements in students' academic success (Barnett, 1998).

Self-Regulation

Heckman, in a series of sophisticated econometric studies, has demonstrated the importance of what he conceptualizes as "noncognitive" factors such as personality traits, persistence, motivation, and charm in determining school and labor market outcomes (Borghans, Duckworth, Heckman, & Weel, 2008; Heckman & Rubinstein, 2001; Heckman, Stixrud, & Urzua, 2006). One critical noncognitive factor for student outcomes is self-regulation, which involves the management of emotions and emotion-related behaviors, focusing attention, planning and problem-solving, and delay of gratification (Barkley, 1997; Casey et al., 2011; Eisenberg et al., 1997; Hoyle & Bradfield, 2010; Kochanska, Murray, & Coy, 1997). Students who show positive self-regulation can demonstrate persistence and attention to tasks and can manage their emotions in a way that allows them to benefit from a collaborative classroom environment (e.g., handling corrective feedback on their performance from a teacher, successfully resolving disagreements with peers). Students who

have difficulty managing negative emotions early on are especially likely to suffer later behavior problems compared with peers (Eisenberg et al., 2005).

Empirical evidence supports a positive relationship between self-regulation and academic outcomes. For example, Fantuzzo and colleagues (2007) found that 8.8 % of variance in early mathematics was uniquely predicted by children's ability to regulate their behavior. McClelland and colleagues (2007) found that not only was behavioral regulation positively related to early academic achievement but that *growth* in behavioral regulation during the course of a preschool year predicted growth in children's school readiness in three areas important to cognitive development.

One special application of self-regulation is the concept of *grit* (Duckworth, Peterson, Matthews, & Kelly, 2007; Duckworth & Quinn, 2009), which is defined as trait-level perseverance and passion for long-term goals, but more commonly described as backbone, chutzpah, or stick-to-it-iveness. In a series of studies, grit was positively associated with undergraduate grade point average, cadet enrollment in a second year at West Point, and better performance in a competitive spelling bee. Grit tends to be associated with older age, higher levels of education, and fewer career changes (Duckworth et al., 2007).

Social and Emotional Learning

Social and emotional learning (SEL; Zins & Elias, 2006) refers to the process of developing the cognitive and emotional capacities to recognize and manage emotions, solve problems effectively, and establish positive relationships with others. SEL involves acquiring and effectively applying the knowledge, attitudes, and skills necessary to allow children to calm themselves when angry, make friends, resolve conflicts respectfully, and make ethical and safe choices (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2007). A growing body of research has demonstrated that SEL promotes positive development among children and youth, reduces problem behaviors, and improves academic performance, citizenship, and health-related behaviors (CASEL, 2007; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

Teacher-Level Predictors of Academic and Social Development

Teachers' high expectations and sense of responsibility for student academic achievement influence positive student outcomes (Goddard, Hoy, & Hoy, 2000; Hinnant, O'Brien, & Ghazarian, 2009; Pedersen, Faucher, & Eaton, 1978) and enable the implementation of student-centered approaches (Bryk & Schneider,

2002; Osher, Sandler, & Nelson, 2001). Teachers who are themselves socially and emotionally competent can encourage prosocial communication among students, create a positive climate (Brackett, Katulak, Kremenitzer, Alster, & Caruso, 2008), and influence student conduct, engagement, connectedness to school, and academic performance (Hawkins, 1997; Schaps, Battistich, & Solomon, 1997, 2004; Sutton & Wheatley, 2003; Wentzel, 2002). Alternatively, teachers whose own social and emotional skills are poorer tend to experience greater stress and burnout (Tsouloupas, Carson, Matthews, Grawitch, & Barber, 2010).

In this section, we examine one well-known teacher correlate of student outcomes (training/experience) and two lesser-known indicators (mindfulness and cultural competence) that are beginning to accrue research evidence for their significance.

Teacher Training

Although researchers consistently demonstrate the overwhelming importance of teachers for student outcomes (e.g., Rivkin, Hanushek, & Kain, 2005), the specific characteristics that make a teacher effective are vigorously debated. Researchers examining measurable proxies of quality, such as certification, academic degrees, and years of experience, have found that these elements bear some relationship to student achievement, but the size of the effect is generally modest (Ballou & Podgursky, 2000; Darling-Hammond, 2000; Darling-Hammond & Youngs, 2002; Goldhaber & Brewer, 1999; U.S. Department of Education, 2002). Alternatively, teachers' high verbal ability and content knowledge may be more strongly related to student academic achievement than specialized training on how to teach (e.g., student teaching, education course work; Ballou & Podgursky, 2000; Finn, 1999).

Mindfulness

In a meta-analysis of more than 100 studies of classroom management, Marzano, Marzano, and Pickering (2003) found that "mental set" had the largest effect on reductions in disruptive behavior ($d = -1.3$). The construct of mental set is similar to Langer's (1997) construct of "mindfulness," which refers to "a heightened sense of situational awareness and a conscious control over one's thoughts and behavior relative to that situation" (Marzano et al., 2003, p. 65). In contrast, "mindlessness" is a state of "automatic pilot" where one operates with little conscious awareness. Mindfulness also includes emotional objectivity. Teachers who remain cool under pressure addressing disciplinary issues in a "matter-of-fact" way without taking behaviors personally are most effective classroom managers (Boles, Dean, Ricks, Short, & Wang, 2000; Jennings & Greenberg, 2009; Kokkinos, 2006; Pines & Keinan, 2005). Research shows that high levels of teacher stress and burnout are

linked to perceived high demands and low control in the job (Betoret, 2009; Santavirta, Solovieva, & Theorell, 2007).

To develop teachers' skills for coping with stress, evidence-based SEL programs like the RULER approach (Recognizing, Understanding, Labeling, Expressing, and Regulating emotions; Brackett et al., 2008) and the Garrison Institute's CARE (Cultivating Awareness and Resilience in Education) provide training for teachers, administrators, and students. Evidence shows that RULER improves academic achievement, reduces problem behaviors, and creates a more positive climate (Reyes, Brackett, Rivers, Elbertson, & Salovey, 2012; Reyes et al., 2012). CARE is a professional development program that reduces stress and promotes improvements in teachers' well-being, motivational orientation/efficacy, and mindfulness, as well as improves teachers' organizational, instructional, and emotional support for students (Jennings, 2011). Evaluations of this program indicate that participating teachers report an improved sense of well-being and the ability to provide effective social and emotional support for students (Jennings, Snowberg, Coccia, & Greenberg, 2011), particularly among teachers working in urban settings.

Cultural Competence

Cultural discontinuity between teachers and students, which includes a lack of understanding of behavioral and learning styles, may contribute to lower expectations as well as disproportionate rates of disciplinary referrals, suspensions, and expulsions (Boykin, 1983; Gay, 2000; Skiba, Michael, Nardo, & Peterson, 2000). These problems may be exacerbated by attribution biases and what social psychologists conceptualize as aversive racism (Artiles, Kozleski, Trent, Osher, & Ortiz, 2010; Dovidio, Gaertner, & Pearson, 2009). Cultural competence, which involves the capacity to "step outside of our own framework" (Harry, 1992, p. 334) and to treat individuals individually while respecting and acknowledging their cultural beliefs and values, has been suggested as key to addressing the discontinuity (Osher et al., 2004; Osher, et al., 2012).

School-Level Predictors of Academic and Social Development

In this section, we briefly review selected building-level factors that influence student development: conditions for learning, connectedness, and discipline.

School Climate and Conditions for Learning

"Conditions for learning" refer to those aspects of school climate that are proximally related to learning and development. There are at least four conditions for learning: (a) physical and emotional safety, (b) challenge and engagement, (c) support for students and their connection to school, and (d) the social and emotional competencies of students. These conditions can be facilitated by student support, positive behavioral approaches, robust curricula, strong pedagogy, and support for social and emotional learning (Lee, Smith, Perry, & Smylie, 1999; Osher, Dwyer, & Jimerson, 2006; Osterman, 2000; Ryan & Patrick, 2001; Thuen & Bru, 2009).

A National Research Council report (Bowman, Donovan, & Burns, 2001) pointed out that "one of the most consistent findings in the early childhood literature is that an emotionally warm and positive approach in learning situations leads to constructive behavior in children." Pianta and colleagues (2008) have shown that better emotional quality of classroom interactions positively predicts growth in reading and math achievement from first through fifth grades. Starting early is very important, but across all years of schooling, enhancing social and emotional behaviors can have a strong impact on success in school and ultimately in life (Zins, Bloodworth, Weissberg, & Walberg, 2004).

Conditions for learning are shaped by several characteristics, such as caring and supportive interpersonal relationships (e.g., consistent acknowledgment of all students, recognition for good work), student and adult social and emotional, approaches to discipline (e.g., school wide positive behavioral supports, restorative justice), opportunities for meaningful participation for students (e.g., student decision-making with classroom management), effective classroom management practices, and high levels of academic and behavioral expectations and support to realize them (Osher, Sidana, & Kelly, 2008).

The literature concerning climate and conditions for learning shows an association between level of school climate and academic achievement (e.g., Kendziora, Osher, & Chinen, 2008; Klem & Connell, 2004; Osher, Spier, Kendziora, & Cai, 2009; Spier, Cai, Osher, & Kendziora, 2007). In addition, researchers have shown that, regardless of the level of school climate, improving school climate is associated with *increases* in student performance in reading, writing, and mathematics (Osher et al., 2009).

School Connectedness

Connectedness denotes students' school experiences and perceptions that learning is important, their sense that teachers are respectful and supportive, and feelings that other students and school personnel are safe, close, and united across the school. In fact, longitudinal evidence suggests that school connectedness is the

second most salient predictor (family connectedness being the first) for protecting adolescents from emotional distress, eating disorders, and suicidal ideation and attempts (Blum, McNeely, & Rinehart, 2002; Cairns & Cairns, 1994; Goodenow, 1993; Resnick et al., 1997; Resnick, Harris, & Blum, 1993).

Connections with adults (teachers, administrators) are strongly associated with positive academic and social outcomes for students (Barber & Olsen, 2004; Bergin & Bergin, 2009; Cornelius-White, 2007; Hamre & Pianta, 2003; Wu, Hughes, & Kwok, 2010), as well as a reduction in dropout and delinquent behaviors (Bernard, 2004; Colarossi & Eccles, 2003; Croninger & Lee, 2001; Hughes, Cavell, & Jackson, 1999; Kesner, 2000; Kirby, 2001; Pianta, Hamre, & Stuhlman, 2003). Students' perceptions that teachers care about them are also associated with their academic success and well-being (Barber & Olsen, 1997; Connell, Halpern-Felsher, Clifford, Crichlow, & Usinger, 1995; Hamre & Pianta, 2005; Wentzel & Wigfield, 1998).

School Discipline Practices

The approaches schools take to ensure compliance with behavioral expectations matter for student outcomes. Punitive approaches to discipline have been linked to a range of negative outcomes (Balfanz & Boccanfuso, 2007; Dishion & Dodge, 2005; Gottfredson, Gottfredson, Payne, & Gottfredson, 2005; Gregory, Skiba, & Noguera, 2010; Raffaele Mendez & Knoff, 2003; Skiba & Rausch, 2004, 2006). For example, after following approximately one million Texas public school students for nearly 6 years, Fabelo and colleagues (2011) found that expulsion from school significantly increased the likelihood that a student would repeat a grade, not graduate, and/or be adjudicated into the juvenile justice system.

Despite the availability of risk assessments to determine an appropriate level of treatment (e.g., expulsion vs. counselor support) for individual delinquent behavior, zero tolerance policies in schools result in the application of sanctions regardless of the context or the student's potential risk to others (Curwin & Mendler, 1999). This is particularly salient for minorities and students with emotional and behavioral disorders who are disproportionately suspended and expelled from schools (Gregory et al., 2010; Morrison et al., 2001; Osher, Morrison, & Bailey, 2003; Osher, Woodruff, & Sims, 2002; Townsend, 2000), thereby exacerbating lost opportunities to learn, as well as increasing disengagement and the risk of dropout. Harsh intervention for youth whose behavior does not warrant intense disciplinary action may also be harmful, particularly without additional support for low-risk youth (Mendez, 2003).

In contrast to reactive and exclusionary punishments, schools are widely adopting models that explicitly teach behavioral expectations. On the frontier in school discipline are ecological approaches focused on fostering self-discipline and self-control (Osher, Bear, Sprague, & Doyle, 2010). These strategies focus on improving the efficacy and holding power of the classroom activities in which

students participate (see Bear, 2010; Doyle, 2006). Early research in this area has shown that proactive strategies for managing classroom group structures can promote student engagement in the classroom (Gump, 1990; Kounin, 1970).

Promotion and Prevention in Schools

As a universal institution, schools are in a unique position both to promote student health and to benefit from the presence of healthy students. Long lists of exemplary school-based promotion and prevention programs have been developed and are widely available (e.g., Center for the Study and Prevention of Violence, 2006). School-based prevention programs are most effective when they are tailored based on the characteristics of participating children and youth, such as age, culture, and community context. Programs for children in elementary schools are most effective when they focus on improving both academic and social/emotional learning, whereas programs for middle school and high school students are most effective when they focus on building academic and social competences, such as good study habits, drug resistance skills, and positive relationships with peers (National Institute on Drug Abuse, 2011). Prevention efforts can be especially effective—even among high-risk youth—when provided at key transition points (such as moving from middle school to high school) and provided to the general population (rather than singling out and labeling individuals; Institute of Medicine, 2009). Although specific violence and substance abuse prevention curricula are available, schools can also take steps to prevent adverse outcomes by ensuring that teachers apply good classroom management practices, such as reinforcing positive behaviors (National Institute on Drug Abuse, 2011) and creating a positive school climate.

In the remainder of this section, we focus on student health, nutrition, and mental health and consider social and emotional learning as an organizing framework for school interventions.

Physical Health

To perform well academically, students must first attend school (Windham, Bohrnstedt, Brown, Seidel, & Kendziora, 2011). It is not surprising that students with chronic conditions such as asthma and obesity have poorer attendance than healthy students and that improving student health also improves student attendance. There is evidence that when schools increase student health conditions for students, such as by facilitating student access to health services and increasing physical activity, student absences decrease (Basch, 2011a). Therefore, schools can take steps to improve student attendance by focusing on improving student health.

Schools have long promoted student health at the population level through physical education classes and recess. Children who are physically active for at

least 60 min per day have a significantly reduced risk of later cardiovascular disease and diabetes relative to less active children (Physical Activity Guidelines Advisory Committee, 2008). In addition, exercise has been demonstrated experimentally to improve academic performance (Stevens, To, Stevenson, & Lochbaum, 2008; Telford et al., 2012).

There is substantial evidence that allocating time for physical activity during the school day does not detract from children's academic performance and may even improve it. In 2010, the Centers for Disease Control and Prevention (CDC) undertook a review of 50 studies that examined the relationship between physical activity time at school and academic performance. Just over half of these studies showed that student participation in school-based physical activities conferred positive benefits on children's academic performance, just under half showed no effect on academic performance, and only one study showed a negative effect (CDC, 2010).

Nutrition

Adequate nutrition is a cornerstone of child health and development. Multiple studies have demonstrated the negative effect of food insecurity on school-age children's academic achievement, behavior, and social relationships with peers (Cook & Frank, 2008). The federally funded National School Lunch Program exists to address food insecurity among students, but even when food is provided to students, it may fail to meet existing nutrition requirements (New America Foundation, 2011). Among undernourished populations, school breakfast programs improve attendance, cognitive functioning, and academic performance (Basch, 2011a, 2011b).

Mental Health

For children with mental health needs, schools, not the specialty mental health sector, are the primary providers of services (Hoagwood & Erwin, 1997). Schools have not been very successful, however, in meeting the needs of children with emotional disturbances. Compared with other students with disabilities, students with emotional disturbances are identified later and are more likely to be in restrictive placements and drop out of school (U.S. Department of Education, 2011; Wagner et al., 2006). Children with mental health issues that do not involve disruptive behavior, such as those with depression or anxiety, are particularly likely to be identified late or not at all. In addition, research suggests that there is a "dual track" to services, in which a disproportionate number of children of color as compared with Caucasian children first receive mental health services only after they encounter the child welfare or juvenile justice system (Huang, 2007).

Social and Emotional Learning Interventions

Over the past two decades, researchers have produced increasingly compelling evidence that interventions promoting SEL improve academic performance. Zins, Weissberg, Wang, and Walberg reported that students who become more self-aware and confident about their learning abilities try harder in school. Students who set high academic goals, have self-discipline, motivate themselves, manage their stress, and organize their approach to work learn more and get better grades (Duckworth & Seligman, 2005; Elliot & Dweck, 2005). Furthermore, new research in neuroscience suggests that SEL programs may improve central executive cognitive functions, such as inhibitory control, planning, and set-shifting, by building greater cognitive-affective regulation in prefrontal areas of the cortex (Riggs, Greenberg, Kusché, & Pentz, 2006).

Students in schools that use an evidence-based SEL curriculum (one that has been scientifically evaluated and found effective) significantly improve in their attitudes toward school, their behaviors, and their academic performance (Durlak et al., 2011). Durlak and colleague's (2011) meta-analysis of 30 studies found that SEL results in improvements in students' achievement test scores—by an average of 11 percentile points over students who are not involved in SEL programming. In a multiyear study, Hawkins, Smith, and Catalano (2004) found that by the time they were adults, students who received SEL interventions in grades 1–6 had an 11 % higher grade point average and significantly greater levels of school commitment and attachment to school at age 18 than did other students. Furthermore, students who received SEL interventions showed a 30 % lower incidence of school behavior problems, a 20 % lower rate of violent delinquency, and a 40 % lower rate of heavy alcohol use by age 18.

In a meta-analysis of 207 studies of SEL programs, Durlak and colleagues (2011) found that following the intervention, students demonstrated enhanced SEL skills, attitudes, and positive social behaviors; demonstrated fewer conduct problems; and had lower levels of emotional distress compared with controls. Furthermore, academic performance was significantly improved, with overall mean effect sizes for test scores and grades of 0.27 and 0.33, respectively. SEL programs with the best outcomes are multiyear in duration, use interactive rather than purely knowledge-based instructional methods, and are integrated into the life of the school rather than implemented as marginal add-ons.

Concluding Remarks: Frontiers with Positive Development and Social Emotional Development

The challenge for schools and communities is to create environments where students are supported in developing the skills and capacities to thrive. Doing this successfully involves leveraging and disseminating information about the

connections between learning and academic performance. For example, in Alaska, where the American Institutes for Research developed a survey to assess the impact of a youth development initiative across 18 school districts, we found that not only were several aspects of school climate and connectedness related to student achievement but positive change in school climate and connectedness was related to significant gains in student scores on statewide achievement tests. Our findings showed that, whether a school started with high or low school climate and connectedness or high or low achievement scores, changing that school's climate and connectedness for the better was associated with increases in student performance in reading, writing, and mathematics (Spier et al., 2007). Similarly, the work we have done in previously troubled schools in the South Bronx suggested that mental health supports, when coupled with family engagement and organizational efficacy, were effective in creating more positive conditions for learning in these schools (Kendziora et al., 2008).

Although research that demonstrates the importance of the social and emotional conditions for learning in schools is accumulating, these factors remain relatively marginalized in the education community. Steps that may be taken to create emotionally safe and supportive schools that promote students' positive social, emotional, and academic learning include (a) adopting SEL as a framework for school improvement, (b) helping districts and states develop the capacity to assess and monitor their social and emotional conditions for learning, and (c) providing schools and communities with effective tools and strategies to improve these conditions. The Association for Supervision and Curriculum Development (ASCD, a professional association for educators) has adopted a Whole Child Initiative that stresses an approach consistent with the interventions we have described. ASCD writes, "Each child, in each school, in each of our communities deserves to be healthy, safe, engaged, supported, and challenged. That's what a whole child approach to learning, teaching, and community engagement really is."

References

Artiles, A. J., Kozleski, E. B., Trent, S. C., Osher, D., & Ortiz, A. (2010). Justifying and explaining disproportionality, 1968–2008: A critique of underlying views of culture. *Exceptional Children, 76*(3), 279–299.

Balfanz, R., & Boccanfuso, C. (2007). *Falling off the path to graduation: Middle grade indicators in an unidentified northeastern city*. Baltimore, MA: Center for Social Organization of Schools.

Ballou, D., & Podgursky, M. (2000). Reforming teacher preparation and licensing: What is the evidence? *Teachers College Record, 102*, 5–27.

Barber, B. K., & Olsen, J. (1997). Socialization in context: Connection, regulation, and autonomy in the family, school, and neighborhood and with peers. *Journal of Adolescent Research, 12*, 287–315.

Barber, B. K., & Olsen, J. A. (2004). Assessing the transitions to middle and high school. *Journal of Adolescent Research, 19*(1), 3–30.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65–94.

Barnett, W. S. (1998). Long term effects on cognitive development and school success. In W. S. Barnett & S. S. Boocock (Eds.), *Early care and education for children in poverty: Promises, programs, and long-term results* (pp. 11–44). Albany, NY: SUNY Press.

Basch, C. E. (2011a). Healthier students are better learners: A missing link in school reforms to close the achievement gap. *Journal of School Health, 81*, 593–598.

Basch, C. E. (2011b). Breakfast and the achievement gap among urban minority youth. *Journal of School Health, 81*, 635–640.

Bear, G. G. (2010). *From school discipline to self-discipline*. New York, NY: Guilford.

Bergin, C., & Bergin, D. (2009). Attachment in the classroom. *Educational Psychology Review, 21* (2), 141–170.

Bernard, B. (2004). *Resiliency: What we have learned*. San Francisco, CA: WestEd.

Betoret, F. D. (2009). Self-efficacy, school resources, job stressors, and burnout among Spanish primary and secondary school teachers: A structural equation approach. *Educational Psychology, 29*(1), 45–68.

Blum, R. W., McNeely, C. A., & Rinehart, P. M. (2002). *Improving the odds: The untapped power of schools to improve the health of teens*. Minneapolis, MN: Center for Adolescent Health and Development, University of Minnesota.

Boles, J. S., Dean, D. H., Ricks, J. M., Short, J. C., & Wang, G. (2000). The dimensionality of the Maslach Burnout Inventory across small business owners and educators. *Journal of Vocational Behavior, 56*, 12–34.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Weel, B. (2008). Symposium on noncognitive skills and their development: The economics and psychology of personality traits. *Journal of Human Resources, 43*, 972–1059.

Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2001). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.

Boykin, A. W. (1983). The academic achievement performance of Afro-American Children. In J. Spence (Ed.), *Achievement and achievement motives* (pp. 324–371). San Francisco, CA: Freeman.

Brackett, M. A., Katulak, N. A., Kremenitzer, J. P., Alster, B., & Caruso, D. R. (2008). Emotionally literate teaching. In M. A. Brackett, J. P. Kremenitzer, M. Maurer, M. D. Carpenter, S. E. Rivers, & N. A. Katulak (Eds.), *Emotional literacy in the classroom: Upper elementary*. Port Chester, NY: National Professional Resources.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York, NY: Russell Sage.

Cairns, R. B., & Cairns, B. D. (1994). *Lifelines and risks: Pathways of youth in our time*. New York, NY: Cambridge University Press.

Casey, B. J., Somerville, L. H., Gotlib, I. H., Aydukc, O., Franklin, N. T., Askren, M. K., et al. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences, 108*, 14998–15003.

Center for the Study and Prevention of Violence. (2006, August). *Blueprints model program descriptions*. Boulder, CO: Author.

Centers for Disease Control and Prevention. (2010). *The association between school-based physical activity, including physical education, and academic performance*. Atlanta, GA: U.S. Department of Health and Human Services.

Colarossi, L. G., & Eccles, J. S. (2003). Differential effects of support providers on adolescents' mental health. *Social Work Research, 27*, 19–30.

Collaborative for Academic, Social, and Emotional Learning. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Author.

Connell, J. P., Halpern-Felsher, B., Clifford, E., Crichlow, W. E., & Usinger, P. (1995). Hanging in there: Behavioral, psychological, and contextual factors affecting whether African American adolescents stay in high school. *Journal of Adolescent Research, 10*, 41–63.

Cook, J. T., & Frank, D. A. (2008). Food security, poverty, and human development in the United States. *Annals of the New York Academy of Sciences, 1136*, 193–209.

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research, 77*(1), 113–143.

Croninger, R. G., & Lee, V. E. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *Teachers College Record, 103*, 548–582.

Curwin, R. L., & Mendler, A. N. (1999). Zero tolerance for zero tolerance. *Phi Delta Kappan, 81*, 119–124.

Darling-Hammond, L. (2000). Reforming teacher preparation and licensing: Debating the evidence. *Teachers College Record, 102*, 28–56.

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Researcher, 31*(9), 13–25.

Dishion, T. J., & Dodge, K. A. (2005). Peer contagion in interventions for children and adolescents: Moving towards an understanding of the ecology and dynamics of change. *Journal of Abnormal Child Psychology, 33*, 395–400.

Dovidio, J. F., Gaertner, S. L., & Pearson, A. (2009). The nature of contemporary prejudice: Insights from aversive racism. *Social and Personality Psychology Compass, 3*, 314–338.

Doyle, W. (2006). Ecological approaches to classroom management. In C. Evertson & C. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 97–125). New York, NY: Lawrence Erlbaum.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*, 1087–1101.

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment, 91*, 166–174.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*, 939–944.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405–432.

Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In C. Ames & R. Ames (Eds.), *Research on motivation in education: Goals and cognitions* (Vol. 3, pp. 139–186). New York, NY: Academic.

Eisenberg, N., Fabes, R., Shepard, S., Murphy, B., Guthrie, I., Jones, P., et al. (1997). Contemporaneous and longitudinal prediction of children's social functioning from regulation and emotionality. *Child Development, 68*, 642–664.

Eisenberg, N., Sadovsky, A., Spinrad, T. L., Fabes, R. A., Lasoya, S. H., Valiente, C., et al. (2005). The relations of problem behavior status to children's negative emotionality, effortful control, and impulsivity: Concurrent relations and prediction of change. *Developmental Psychology, 41*, 193–211.

Elliot, A. J., & Dweck, C. S. (Eds.). (2005). *Handbook of competence and motivation*. New York, NY: Guilford Press.

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., III, & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. New York, NY: Council of State Governments Justice Center.

Fantuzzo, J., Bulotsky-Shearer, R., McDermott, P. A., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology Review, 36*, 44–62.

Finn, C. E. (1999). Foreword. In M. Kanstoroom & C. E. Finn Jr. (Eds.), *Better teacher, better schools* (pp. v–vii). Washington, DC: Fordham Foundation.

Gay, G. (2000). *Culturally responsive teaching: Theory, research and practice*. New York, NY: Teachers College Press.

Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal, 37*, 479–507.

Goldhaber, D., & Brewer, D. (1999). Teacher licensing and student achievement. In M. Kanstoroom & C. E. Finn Jr. (Eds.), *Better teachers, better schools* (pp. 83–102). Washington, DC: Fordham Foundation.

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the Schools, 30*, 79–90. doi:10.1002/1520-6807(199301)30:1<79:: AID-PITS2310300113>3.0.CO; 2-X.

Gottfredson, G. D., Gottfredson, D. C., Payne, A. A., & Gottfredson, N. C. (2005). School climate predictors of school disorder: Results from a national study of delinquency prevention in schools. *Journal of Research in Crime and Delinquency, 42*, 412–444.

Gregory, A., Skiba, R. J., & Noguera, P. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher, 39*, 59–68.

Gump, P. V. (1990). A short history of the Midwest Psychological Field Station. *Environment and Behavior, 22*, 436–457.

Hamre, B. K., & Pianta, R. C. (2003). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*, 625–638.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*, 949–967.

Harry, B. (1992). *Cultural diversity, families, and the special education system: Communication and empowerment*. New York, NY: Teachers College Press.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MA: Paul H. Brookes.

Hawkins, J. D. (1997). Academic performance and school success: Sources and consequences. In R. P. Weissberg, T. P. Gullotta, R. L. Hampton, B. A. Ryan, & G. R. Adams (Eds.), *Healthy children 2010: Enhancing children's wellness* (pp. 278–305). Thousand Oaks, CA: Sage.

Hawkins, J. D., Smith, B. H., & Catalano, R. F. (2004). Social development and social and emotional learning. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning. What does the research say?* (pp. 135–150). New York, NY: Teachers College Press.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91*, 145–149.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*, 411–482.

Hinnant, J. B., O'Brien, M., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of Educational Psychology, 101*, 662–670.

Hoagwood, K., & Erwin, H. (1997). Effectiveness of school-based mental health services for children: A 10-year research review. *Journal of Child and Family Studies, 6*, 435–451.

Hoyle, R. H., & Bradfield, E. K. (2010, February). *Measurement and modeling of self-regulation: Is standardization a reasonable goal?* Working paper prepared for the National Research Council Workshop on Advancing Social Science Theory: The Importance of Common Metrics, Washington, DC.

Huang, L. (2007, May). *"Wicked problems" and collective solutions: Reducing disparities in mental health care for underserved populations*. Plenary presentation at Building Family Strengths Conference of the Research and Training Center on Family Support and Children's Mental Health, Portland, OR.

Hughes, J. N., Cavell, T. A., & Jackson, T. (1999). Influence of teacher-student relationship on childhood conduct problems: A prospective study. *Journal of Clinical Child Psychology, 28*, 173–184.

Institute of Medicine. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academy Press.

Jennings, P. A. (2011). Promoting teachers' social and emotional competencies to support performance and reduce burnout. In A. Cohan & A. Honigsfeld (Eds.), *Breaking the mold of preservice and inservice teacher education: Innovative and successful practices for the twenty-first century*. New York, NY: Rowman & Littlefield.

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research, 79*, 491–525.

Jennings, P. A., Snowberg, K. E., Coccia, M. A., & Greenberg, M. T. (2011). Improving classroom learning environments by Cultivating Awareness and Resilience in Education (CARE): Results of two pilot studies. *Journal of Classroom Interaction, 46*(1), 37–48.

Kellam, S. G., Branch, J. D., Agrawal, K. C., & Ensminger, M. E. (1975). *Mental health and going to school: The Woodlawn program of assessment, early intervention, and evaluation*. Chicago, IL: University of Chicago Press.

Kendziora, K., Osher, D., & Chinen, M. (2008). *Student connection research: Final narrative report to the Spencer Foundation*. Washington, DC: American Institutes for Research.

Kesner, J. E. (2000). Teacher characteristics and the quality of child-teacher relationships. *Journal of School Psychology, 28*(2), 133–149.

Kidron, Y., & Osher, D. (2012). The history and direction of research on prosocial education. In P. Brown, M. Corrigan, & A. H. D'Allesandro (Eds.), *The handbook of prosocial education*. Lanham, MD: Roman & Littlefield.

Kirby, D. (2001). *Emerging answers: Research findings on programs to prevent teen pregnancy*. Washington, DC: National Campaign to Prevent Teen Pregnancy.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health, 74*, 262–273.

Kochanska, G., Murray, K., & Coy, K. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*, 263–277.

Kokkinos, C. M. (2006). Factor structure and psychometric properties of the Maslach Burnout Inventory: Educators survey among elementary and secondary school teachers in Cyprus. *Stress and Health, 22*, 25–33.

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York, NY: Holt, Rinehart & Winston.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*, 178–181.

Langer, E. J. (1997). *The power of mindful learning*. New York, NY: Merloyd Lawrence.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

Lee, V. E., Smith, J. B., Perry, T. E., & Smylie, M. A. (1999). *Social support, academic press, and student achievement: A view from the middle grades in Chicago*. Chicago, IL: Consortium on Chicago School Research.

Marzano, R. J., Marzano, J. S., & Pickering, D. J. (2003). *Classroom management that works*. Alexandria, VA: ASCD.

McClelland, M. M., Connor, C. M., Jewkes, A. M., Cameron, C. E., Farris, C. L., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43*, 947–959.

Mendez, L. M. R. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. In J. Wald & D. Losen (Eds.), *New directions for youth development: Deconstructing the school-to-prison pipeline* (pp. 17–34). Hoboken, NJ: Wiley & Sons.

Morrison, G. M., Anthony, S., Storino, M. H., Cheng, J. J., Furlong, M. J., & Morrison, R. L. (2001). School expulsion as a process and an event: Before and after effects on children at-risk for school discipline. *New Directions for Youth Development, 92*, 45–71.

National Institute on Drug Abuse. (2011). *Info facts: Lessons from prevention research*. Retrieved January 27, 2012, from http://www.drugabuse.gov/publications/infofacts/lessons-prevention-research

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: Bradford Books.

New America Foundation. (2011). *Background and analysis: School nutrition programs*. Retrieved January 25, 2012, from http://febp.newamerica.net/background-analysis/federal-school-nutrition-programs

Osher, D., Bear, G. G., Sprague, J. R., & Doyle, W. (2010). How can we improve school discipline? *Educational Researcher, 39*, 48–58.

Osher, D., Cartledge, G., Oswald, D., Sutherland, K. S., Artiles, A. J., & Coutinho, M. (2004). Issues of cultural and linguistic competency and disproportionate representation. In R. B. Rutherford, M. M. Quinn, & S. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 54–77). New York, NY: Guilford Press.

Osher, D., Coggshall, J., Colombi, G., Woodruff, D., Francois, S., & Osher, T. W. (2012). Building school and teacher capacity to eliminate the school-to-prison pipeline. *Teacher Education and Special Education, 35*, 284–295.

Osher, D., Dwyer, K., & Jimerson, S. (2006). Foundations of school violence and safety. In S. Jimerson & M. Furlong (Eds.), *Handbook of school violence and school safety: From research to practice* (pp. 51–71). Mahwah, NJ: Lawrence Erlbaum Associates.

Osher, D., Morrison, G., & Bailey, W. (2003). Exploring the relationship between student mobility and dropout among students with emotional and behavioral disorders. *Journal of Negro Education, 72*, 79–96.

Osher, D., Sandler, S., & Nelson, C. (2001). The best approach to safety is to fix schools and support children and staff. *New Directions in Youth Development, 92*, 127–154.

Osher, D., Sidana, A., & Kelly, P. (2008). *Improving conditions for learning for youth who are neglected or delinquent*. Washington, DC: National Evaluation and Technical Assistance Center.

Osher, D., Spier, E., Kendziora, K., & Cai, C. (2009). *Improving academic achievement through improving school climate and student connectedness*. Paper presented at the annual meeting of the American Educational Researchers Association, San Diego, CA.

Osher, D., Woodruff, D., & Sims, A. (2002). Schools make a difference: The relationship between education services for African American children and youth and their overrepresentation in the juvenile justice system. In D. Losen (Ed.), *Minority issues in special education* (pp. 93–116). Cambridge, MA: Harvard Education Publishing Group.

Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research, 70*, 323–367.

Pedersen, E., Faucher, T. A., & Eaton, W. W. (1978). A new perspective on the effects of first-grade teachers on children's subsequent adult status. *Harvard Educational Review, 48*, 1–31.

Physical Activity Guidelines Advisory Committee. (2008). *Physical Activity Guidelines Advisory Committee report*. Washington, DC: U.S. Department of Health and Human Services.

Pianta, R., Belsky, J., Vandergrift, N., Houts, R., Morrison, F., & The National Institute of Child Health and Human Development Early Child Care Research Network. (2008). Classroom effects on children's achievement trajectories in elementary school. *American Educational Research Journal, 45*, 365–397.

Pianta, R. C., Hamre, B. K., & Stuhlman, M. W. (2003). Relationships between teachers and children. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7, pp. 199–234). New York, NY: Wiley.

Pines, A. M., & Keinan, G. (2005). Stress and burnout: The significant difference. *Personality and Individual Differences, 39*, 625–635.

Raffaele Mendez, L. M., & Knoff, H. M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education & Treatment of Children, 26*, 30–51.

Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., et al. (1997). Protecting adolescents from harm: Findings from the National Longitudinal Study on Adolescent Health. *Journal of the American Medical Association, 278*, 823–832.

Resnick, M. D., Harris, L. J., & Blum, R. W. (1993). The impact of caring and connectedness on adolescent health and well-being. *Journal of Pediatrics & Child Health, 29*(Suppl 1), S3–S9.

Reyes, M. R., Brackett, M. A., Rivers, S. E., Elbertson, N., & Salovey, P. (2012). The interaction effects of program training, dosage, and implementation quality on targeted student outcomes for the RULER approach to social and emotional learning. *School Psychology Review, 41*(1), 82–99.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., Mashburn, A. J., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*(3), 700–712.

Riggs, N. R., Greenberg, M. T., Kusché, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS curriculum. *Prevention Science, 7*, 91–102. doi:10.1007/s11121-005-0022-1.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*, 417–458.

Rury, J. L. (2002). *Education and social change: Themes in the history of American schooling*. Mahwah, NJ: Erlbaum.

Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal, 38*, 437–460.

Santavirta, N., Solovieva, S., & Theorell, T. (2007). The association between job strain and emotional exhaustion in a cohort of 1028 Finnish teachers. *British Journal of Educational Psychology, 77*, 213–228.

Sattler, J. M., & Dumont, R. (2004). *Assessment of children: WISC-IV and WPPSI-III supplement*. San Diego, CA: Jerome M. Sattler Publisher.

Schaps, E., Battistich, V., & Solomon, D. (1997). School as a caring community: A key to character education. In A. Molnar (Ed.), *The construction of children's character, part II, ninety-sixth yearbook of the National Society for the Study of Education* (pp. 127–139). Chicago, IL: University of Chicago Press.

Schaps, E., Battistich, V., & Solomon, D. (2004). Community in school as key to student growth: Findings from the Child Development Project. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 189–205). New York, NY: Teachers College Press.

Shonkoff, J., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Skiba, R. J., Michael, R., Nardo, A., & Peterson, R. (2000). *The color of discipline: Gender and racial disparities in school punishment*. Bloomington, IN: Indiana Education Policy Center.

Skiba, R., & Rausch, M. K. (2004). *The relationship between achievement, discipline, and race: An analysis of factors predicting ISTEP scores*. Bloomington, IN: Center for Evaluation and Education Policy.

Skiba, R., & Rausch, M. K. (2006). Zero tolerance, suspension, and expulsion: Questions of equity and effectiveness. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 1063–1089). Mahwah, NJ: Lawrence Erlbaum Associates.

Spier, E., Cai, C., Osher, D., & Kendziora, K. (2007, September). *School climate and connectedness and student achievement in 11 Alaska school districts*. Unpublished report, American Institutes for Research, Washington, DC.

Stevens, T. A., To, Y., Stevenson, S. J., & Lochbaum, M. R. (2008). The importance of physical activity and physical education in the prediction of academic achievement. *Journal of Sport Behavior, 31*, 368–388.

Sutton, R. E., & Wheatley, K. F. (2003). Teachers' emotions and teaching: A review of the literature and directions for future research. *Educational Psychology, 15*(4), 327–358.

Telford, R. D., Cunningham, R. B., Fitzgerald, R., Olive, L. S., Prosser, L., Jiang, X., et al. (2012). Physical education, obesity, and academic achievement: A 2-year longitudinal investigation of Australian elementary school children. *American Journal of Public Health, 102*.

Thuen, E., & Bru, E. (2009). Are changes in students' perceptions of the learning environment related to changes in emotional and behavioural problems? *School Psychology International, 30*, 115–136.

Townsend, B. L. (2000). The disproportionate discipline of African-American learners: Reducing school suspensions and expulsions. *Exceptional Children, 66*, 381–391.

Tsouloupas, C. N., Carson, R. L., Matthews, R., Grawitch, M., & Barber, L. K. (2010). Exploring the association between teachers' perceived student misbehaviour and emotional exhaustion: The importance of teacher efficacy beliefs and emotional regulation. *Educational Psychology, 30*, 173–189.

U.S. Department of Education. (2002). *The Secretary's report on teacher quality*. Washington, DC: Author.

U.S. Department of Education. (2011). *30th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2008*. Washington, DC: U.S. Government Printing Office.

Wagner, M., Friend, M., Bursuck, W. D., Kutash, K., Duchnowski, A. J., Sumi, W. C., et al. (2006). Educating students with emotional disturbances: A national perspective on school programs and services. *Journal of Emotional and Behavioral Disorders, 14*, 12–30.

Watkins, M. W., Lei, P.-W., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence, 35*, 59–68.

Wentzel, K. R. (2002). Are good teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development, 73*, 287–301.

Wentzel, K. R., & Wigfield, A. (1998). Academic and social motivational influences on students' academic performance. *Educational Psychology Review, 10*, 155–175.

Windham, A., Bohrnstedt, G., Brown, L., Seidel, D., & Kendziora, K. (2011, December). *Weather, transportation, attendance, and student achievement in the Syracuse City School District*. Washington DC: American Institutes for Research.

Wirkala, C., & Kuhn, D. (2011). Problem-based learning in K–12 education: Is it effective and how does it achieve its effects? *American Educational Research Journal, 48*, 1157–1186. doi:10.3102/0002831211419491.

Wu, J., Hughes, J. N., & Kwok, O. (2010). Teacher–student relationship quality type in elementary grades: Effects on trajectories for achievement and engagement. *Journal of School Psychology, 48*, 357–387.

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2004). The scientific base linking social and emotional learning to school success. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 3–22). New York, NY: Teachers College, Columbia University.

Zins, J. E., & Elias, M. E. (2006). Social and emotional learning. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: Development, prevention, and intervention* (pp. 1–13). Bethesda, MD: National Association of School Psychologists.

Chapter 8
Peer Networks

Michael Mason

Introduction

Developmental science has long established that children play a role in each other's psychological and social development. Peers contribute uniquely, and independently from family factors, in the socialization process. In fact peer relations can be considered one of the primary engines of development for children. As children turn into adolescents, networks are formed by interlocking relationships that evolve overtime and have bearing on mental and physical health, mortality, and well-being (Hartup, 2009). Understanding the function of peer networks on increasing risk for and protection against adverse outcomes has become foundational for prevention scientists. Social networks can be defined as relational ties or linkages between a finite number of people and entities. These links, patterns, and pathways have implications and can be structurally mapped for visual and statistical analysis. However, a social network is more than a structural map, and it can be considered a window into ones' values, meaning, beliefs, status, behavioral and health practices, and social roles. It is a complex, dynamic, and interactive system that represents an ongoing social force in people's lives. Social networks matter and have implications for various outcomes. For example, extensive evidence exists to support the relationship between social networks and health outcomes, such as health status, health behaviors, and health decision-making (Cattell, 2001; Heaney & Isreal, 1997; House, Umberson, & Landis, 1988; Valente, 2003). Recent research has thrust social networks into the popular media with large, longitudinal studies demonstrating the far-reaching influence of distal (three degrees of separation—the friend of your friend's friend) network effects on various important issues such as smoking, alcohol use, depression, happiness, and obesity with adults (Christakis & Fowler, 2009). Given these and other important findings regarding the potency of

M. Mason (✉)
Commonwealth Institute on Child and Family Studies, Virginia Commonwealth University, Richmond, VA, USA
e-mail: mjmason@vcu.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_8,

networks, the need exists to increase understanding, training, and application of peer network research within prevention science.

Because adolescence and young adulthood are critical periods that are associated with elevated base rates for risk behaviors, the focus of this chapter will be on peer networks of adolescents and young adults. This chapter will begin with a brief history of network research, its interdisciplinary roots, and the development of modern network research approaches. Next, conceptual issues surrounding the relational framework of network research is reviewed, which include developmental considerations and pertinent social theories related to peer networks. Methodological issues are then reviewed covering measurement and sampling approaches. Current issues are examined next covering selection and influence, behavioral genetic approaches, and the role of ethnicity. Contextual explanations and approaches are then reviewed. Finally, future directions for peer network research, with an emphasis on technological implications, are examined.

History of Social Network Research

Social network research is inherently interdisciplinary. Network researchers have often identified the work of psychiatrist Moreno (1934) as the foundation upon which social network research and analysis is built. Moreno's research in developing sociometry, the display and study of relationships within groups, provided a foundation for the structural study of relationships common today. Following Moreno were anthropologists Barnes (1954) and Mitchell (1969) who identified patterns of social relations that explained social problems in their field work and first used the term social network to describe "a set of ties linking members of a social system" (Barnes, 1969, p. 54). Social psychologists Bavelas (1948) and Leavitt (1951) studied group processes through structural analysis such as individual positions within groups, identifying constructs such as centrality (prominence within a network), for example. Bronfenbrenner (1943), whose social ecology theory is fundamental to prevention science, began studying reciprocity of positive affect, structural balance, and transitivity (if the relation holds between a first element and a second and between the second element and a third, it holds between the first and third elements) which proved essential to his dynamic and interactive model of development. Sociologists lay claim to the study of networks through their long-standing tradition of analyzing social structures (Freeman, 1977; Wellman & Berkowitz, 1988). The fields of mathematics and statistics contributed through quantifying networks with graph theory and statistical modeling (Borgatti, 1988; Cartwright & Harary, 1956; Katz & Powell, 1955; Wasserman, 1977) leveraging their considerable analytical developments to the further study of social networks. The discipline of public health emerged as leaders in using social network design and analysis to study the transmission and prevention of sexually transmitted diseases, in particular HIV/AIDS (Fisher & Miscovich, 1990; Latkin et al., 1995; Mandell, Latkin, Oziemkowska, Vlahov, & Celentano, 1994).

While not a linear trajectory, the birth and development of social network research have progressed from psychiatric tool that assisted in understanding relationships of the mentally ill to the anthropological, sociological, and social psychological applications on cultures, groups, and group behaviors. These important foundational advances were then leveraged by developmental scientists who integrated many of these findings to create the social ecological model widely used in prevention science. The mathematical, statistical, and public health approaches further advanced network research with powerful analytical tools applied to critical health issues. In review, the history of network research reveals a field that is highly dynamic and interdisciplinary, and thus productive as well as hard to categorize or define in simple terms.

Conceptualizing Peer Networks

Sociometric approaches in social networks have made great strides in increasing the understanding of individuals embedded within relations (webs of connections) and how positioning of individuals within networks has both individual and group effects. Sociometric positions are graphically displayed and analyzed creating powerful visualization of social networks that illustrate social position within defined populations such as school settings. Peer relations that constitute networks can be conceptualized in terms of popularity (the measure of a person being nominated as a friend), or isolation, centrality (most central or connected person in a network), density of friendships (the degree that people in the network know each other), and in many other ways. The ability to use the network as a natural context from which psychosocial functioning is inferred via sociometric status has made significant contributions to developmental science. These advances in understanding structural characteristics of networks have provided the initial linkages between network characteristics and risk and protective factors. More recently, network researchers have pushed their research beyond global categories such as popular and isolated toward more nuanced microexaminations of varying behavioral and cognitive correlates with sociometric categories (Gifford-Smith & Brownell, 2003). For example, the examination of rejected children and proactive and reactive aggression with moderating variables such as gender and age have shed light on unique and specific preventive opportunities within a network framework for at-risk children and youth.

A social network research approach uses a relational framework in addressing risk behaviors, focusing on the multiple interactions that are embedded in proximal and distal systems that the developing person encounters. Among these interactive systems is the peer network, which has potent and malleable effects on the trajectories of developing youth. Peer networks are particularly salient for youth. Navigating and negotiating social networks are critical social developmental tasks for adolescents and young adults that have long-term implications for the developing person. Extensive research has shown that peer context is a very robust predictor of

outcomes such as (a) tobacco, alcohol, and drug use (Bauman & Ennett, 1996; Hawkins, Catalano, & Miller, 1992; Urberg, Degirmencioglu, & Pilgrim, 1997; Valente, Unger, & Johnson, 2005); (b) delinquent behavior (Dishion, Eddy, Haas, Li, & Spracklen, 1997); (c) high-risk sexual behavior (Patterson, Dishion, & Yoerger, 2000); (d) school problems (Cho, Hallfors, & Sanchez, 2005); (e) aggression (Faris & Ennett, 2012); (f) mental health problems (Mason, 2009; Ueno, 2005; van Zalk, Kerr, Branje, Stattin, & Meeus, 2010); and (g) being overweight (Valente, Fujimoto, Chou, & Spruijt-Metz, 2009). Because a relational model is fitting for studying youth and due to the compelling research on the effects of social networks on health outcomes, peer network research has emerged as a critical area of research within prevention science practice.

Much less evidence is available for prosocial effects of social networks, and this may represent a bias in studying negative outcomes or risk factors rather than protective factors in prevention research. Nevertheless, prosocial group involvement has been shown to be predictive of emotional well-being across childhood and adolescence (Pedersen, Vitaro, Barker, & Borge, 2007; Wentzel, Barry, & Cadwell, 2004). Mechanisms such as peer acceptance, belonging, and providing a "secure base" to explore new social situations are associated with emotional well-being (Birch & Ladd, 1996; Vitaro, Boivin, & Bukowski, 2009). Microsocial process research has found that peer prosocial behaviors stimulate or activate internal motivation of index participants (Barry & Wentzel, 2006) as well as through "conformity training" where peers approve of each other's prosocial behaviors and disapprove of antisocial behaviors (Wentzel, Filisetti, & Looney, 2007). Social groups can create contexts of intimacy, emotional security, affection, mutual support (Vitaro et al., 2009), companionship and social support (Birch & Ladd, 1996), and social visibility (Berndt, 2002). Supportive friendship has been studied as a moderator or protection against psychological and behavioral problems often associated with peer rejection (Lansford et al., 2007) or with negative experiences within families (Bolger, Patterson, & Kupersmidt, 1998).

Developmental Contingencies

The study of peer networks has a strong foundation of interdisciplinary research undergirded by developmental science. For example, understanding the developmental variances of peer relations, behaviors, and interactions among developing youth is critical for social network-based interventions. As children age, their peer relations change in quality and quantity, relative to their cognitive, physical, and social development. In general, peer influence accelerates in early adolescence as time with peers increases, more importance is associated with peer approval and advice, and thus, this serves as a primary source of identity formation (Brown & Klute, 2003; Henrich, Kuperminc, Sack, Blatt, & Leadbeater, 2000; Molloy, Gest, & Rulison, 2011). Children's social groups typically range in size from 3 to 10 and are generally composed of the same sex (Brown & Klute, 2003). Children transition

through middle school to high school and enter adolescence, representing new opportunities for risk and health promotion. Typically, by early adolescence, cliques become increasingly heterosexual (Brown, 1990). This change in gender composition is considered a mechanism to secure one's place in a larger and more diverse setting such as middle school and high school, where more diversity of friends can provide more support. Network density (the degree that people in a network know each other) may reduce children's vulnerability to loneliness by creating a greater sense of belonging and stronger feelings of security compared to less dense networks where children are pulled in multiple directions (Parker, Rubin, Erath, Wojslawowicz, & Buskirk, 2006). Many teens transition into college settings where they join cohorts that are at elevated risk for substance abuse and mental disorders. Throughout these developmental transitions, social networks change and selection and influence of peers exert differing effects. A number of developmentally based factors have moderating effects on peer influence within group settings, such as age (Stattin, Gustafson, & Magnusson, 1989), stability of friendships (Berndt, Hawkins & Jiao, 1999), self-reported susceptibility to peer influence (Schulenberg et al., 1999), context (Mason et al., 2010), prior externalizing problems (Gardner, Roth, & Brooks-Gunn, 2009), parenting style (Mounts & Steinberg, 1995), refusal assertiveness (Epstein & Botvin, 2002), close friend's behavior (Glaser, Shelton, & van den Bree, 2010), peer rejection (Dishion, Patterson, Stoolmiller, & Skinner, 1991), levels of self-regulation (Piehler & Dishion, 2007), and group identification (Kiesner, Cadinu, Poulin, & Bucci, 2002). Peer influence then is considered complex and typically indirect in the effects on individuals. Moderating variables appear to be both malleable (externalizing problems, group identification, self-regulation, and susceptibility) and unmalleable (age, context, parenting style, friend's behavior) and therefore create a complex interactive model, with differing possibilities of preventive interventions. Limited research is available for adult network influence, primarily focused on work settings which implicate the occupational network and associated norms that influence heavy alcohol consumption (Trice & Sonnfnstuhr, 1988). One study found that prior to marriage, the networks of heavy-drinking adults did not differ for emotional, financial, and practical support compared to regular or infrequent drinkers and that the composition of the networks significantly changed after marriage for both men and women (Leonard, Kearns, & Mudar, 2000). Thus, prevention interventions must be adapted to the fluid nature of social networks across these developmental periods.

Social Theories of Peer Effects

A brief summary of the prominent social theories provides an overview for understanding influence and selection within the peer context. Symbolic interaction is developmental interpersonal theory that posits children experience themselves indirectly through the responses of other members of their significant peer groups

(Mead, 1934). Closely related to this self-reflective theory is imitation or modeling of peer behaviors as a mechanism to adapt to social norms and advance in the socialization processes. Bandura's (1996) social cognitive theory of human development asserts that peer influence operates through behavioral modeling or observational learning. However, this specific form of modeling is not mere mimicry; rather, it conveys complex sociocultural rules for generative and innovative behavior and thus also aids in the socialization processes.

Adolescents also have direct experience vicariously by observing peers' behaviors and related consequences. Dishion et al. (1997) advanced a peer contagion theory named deviance training and have studied this among adolescents. Deviance training is a process defined as the interpersonal exchanges that occur between peers that promote deviant behaviors coupled with positive affect, such as laughter. The critical mechanism is eliciting positive affect, through laughter, about past or future possible deviant actions. Deviant training between friends predicted delinquency (Dishion, Spracklen, Andrews, & Patterson, 1996), violence (Dishion et al., 1997), and substance use and high-risk sexual behavior (Patterson et al., 2000). The harmful effects of interventions that aggregate deviant youth have come to the attention of prevention researchers, largely based upon Dishion and Andrews' (1995) study that demonstrated iatrogenic effects over a 3-year period. Not all interventions that group similar peers (selected level) produce harmful effects, and the effects appear to vary by developmental and contextual differences. When interventions target problem behaviors of high interest to group members (e.g., sex and drug use for adolescents), these targeted behaviors appear most vulnerable to deviant peer contagion. The location of the intervention also appears to play a role in the contagion processes. Interventions conducted in settings that are removed from familiar, high-risk settings are less likely to have iatrogenic effects compared to interventions conducted in high-risk settings (Dishion, Dodge, & Lansford, 2006).

Other related explanatory theories are differential association theory—learning about the positives of drug use through their friends' favorable attitudes toward drug use (Sutherland & Cressey, 1978)—and the theory of reasoned action which is the perceptions of (peer) social norms and one's expectations regarding the cost and benefits of engaging in the behavior, which influences one's own intentions to act (Fishbein & Ajzen, 1981). Peer selection effects can be understood through social comparison theory (Festinger, 1954) that asserts individuals to compare themselves in order to formulate and evaluate their opinions and abilities, thereby providing motivation to affiliate with similar peers. These theories of the dynamic processes of modeling, sharing values and norms, social learning, and cost–benefit analyses illuminate the complexity of influence and selection effects on youth behaviors.

Methodological Issues

Measuring Sociometric Status

The great majority of sociometric analysis is conducted within school settings which allow access to defined populations to construct complete network sociometric status graphs. Youth are typically asked to nominate from a roster of their peers (classmates) whom they like the most, or play with the most, and whom they dislike the most, or are least likely to play with. This method allows researchers to identify popular, rejected, central, and other categorical types of children. Peer ratings are similar but use a scale reflecting degree of each category, e.g., 5—like very much, 1—dislike very much, etc. This method is much more time intensive and is particularly useful in smaller settings, e.g., moderately sized classrooms (Gifford-Smith & Brownell, 2003). Another common method of constructing networks is the social cognitive map procedure (SCM) in which students are asked to identify students who "hang out or spend a lot of time together at school." The results produce multiple subgroups that can be analyzed for centrality (nomination frequency). Individuals can be examined for their positions within subgroups as well as the place of the subgroup within the classroom, grade level, or school. These subgroups can then be analyzed by size, interconnectedness to other groups, stability, and, with longitudinal data, formation and change within these groups overtime.

Egocentric/Personal Networks

Networks constructed from the focal actor or ego allows researchers to examine networks based upon ego-nominated close ties. Individuals are asked to generate a list of people (alters) that they typically associate with, spend time with, discuss important issues with, or receive social support from, for example. Next, the ego is asked a series of questions about each alter based upon the research focus. Personal networks are then constructed and analyzed in terms of characteristics of the alters. Questions targeting risk and protective factors are often asked about each alter, producing risk-protection ratios or other values associated with the overall network characteristics (Marsden, 1990). Because egocentric networks are personally defined by the youth, varying levels of subjective importance of the ego's network as well as with specific alters can be accessed and utilized in analyzing these data. For example, group identification, defined as the feelings of involvement and pride that an individual feels toward a group, has been shown to have predictive capability in understanding network effects, with group identification moderating the effects of peer group on individual's change in delinquent behavior over a 1-year period (Kiesner et al., 2002). Thus, subjectively formed networks, which can be considered more natural or realistic than complete classroom or school-derived networks, have an important role in prevention research.

Respondent-Driven Sampling

Finally, a contextually based social network sampling and analytical methodology is respondent-driven sampling (RDS), sometimes referred to as snowball sampling. RDS has developed into a sophisticated network methodology and provides a middle ground compliment to egocentric and saturated methods of social network analysis (Wejnert, 2010). RDS provides data on individuals who are connected through a common behavior (e.g., drug use), thereby allowing for a structural analysis. A strength of this method is that it samples networks within subjects' social–contextual lives, strengthening the external validity of the data collected.

Software Development

The birth and growth of microcomputing led to significant advances in the development of social network analysis software. Many of the popular programs were developed as open-source software, allowing the growing research community to participate by providing feedback for program improvement. A widely used program for network analysis is Pajek (Batagelj & Mrvar, 1998), as is UCINET (Borgatti, Everett, & Freeman, 1992). Another popular program is SIENA (Snijders, Steglich, Schweinberger, & Huisman, 2007) which enables the analysis of network data with advanced features for longitudinal and simulation data analysis features. All of these programs have contributed to the advancement of peer network research, allowing complicated analysis to mirror increasingly complex research questions.

Social network methodology continues to develop and expand with new applications for social relational research. No definitive guide of which method is best suited for which investigation. Context will drive much of these decisions, as network research expands beyond the school setting, different measures and methods will be utilized across various contexts. Direct comparison studies where methods are varied within a large investigation are recommended to determine which method best fits which question (Cillessen, 2009; Maassen, van der Linden, Goosens, & Bokhorst, 2000; Terry & Coie, 1991).

Current Issues

In this section issues are reviewed that have had and are having significant influence on study of peer networks. This section begins by examining the important methodological and interpretative issue of selection and influence. Next, the role and insights of current behavioral genetic approaches are reviewed. This rapidly growing area of study is providing new understandings of development and adjustment

and is challenging long-held positions within the field of peer networks. The roles of ethnicity and sex differences are examined adding more detail to the complexity of peer network research. Finally, contextual approaches toward peer networks are reviewed with a critique of the historical decontextualized understanding of networks.

Selection and Influence

A long-standing issue in the social network community concerns the distinction between homophily (i.e., peer selection), or the tendency of people with similar characteristics (e.g., age, race, religion, behaviors, occupation, etc.) to form social bonds (Byrne, 1971; Lazarfeld & Merton, 1954; McPherson, Smith-Lovin, & Cook, 2001), and influence, where a person's behavior is influenced, to some degree, by that person's social contacts (Friedkin, 1998; Oetting & Donnermeyer, 1998). An important issue for prevention science is determining selection and influence effects relative to causation. This task has implications for study designs, such as the timing for prevention interventions. Several studies provide support for the role of both influence and selection mechanisms on behaviors such as substance use (Bauman & Ennett, 1996; Cottrell, 2007; Hussong, 2002; Kirke, 2006; Wills & Cleary, 1999). Recent efforts to better distinguish between peer selection and influence effects in longitudinal, complete sociometric data sets have employed cutting-edge statistical and simulation models of network dynamics to allow for inferences regarding the role of peer selection versus influence (Steglich, Sniijders, & Pearson, 2010). In presenting a new statistical methodology which appears to be the most rigorous to date, to differentiate selection and influence effects on substance use, Steglich and colleagues used a school sample of 160 students, aged 12–13 years old, focusing on smoking and alcohol use. While selection effects were significant, Steglich and colleagues found that influence had stronger effects than selection for both smoking and alcohol use. While these youth selected peers based upon similar behaviors, the peer network had greater influence effects overtime. In an earlier study using the same sample but including cannabis as an outcome, Pearson, Steglich, and Snijders (2006) found students were more likely to assimilate their behavior to match that of their friends with regard to drinking and cannabis use but not with regard to smoking behavior. Selection effects were more prominent among drinkers and smokers and influence effects were more prominent among cannabis users. As interaction increases with antisocial peers, children's adoption and compliance with social norms in community and school settings are disrupted and can have long-lasting consequences (Dishion & Piehler, 2009).

Peer Attitudes and Substance Use

Peer attitudes are potent predictors of substance use; the effects of these attitudes vary by gender and race/ethnicity and by type of substance. In a recent large-scale, national study, Mason, Mennis, Linker, Bares, and Zaharikas (In Press) examined the relationship between adolescents' perceptions of their close friends' attitudes about substance use and their own use of cigarettes, alcohol, and marijuana. Across all racial and gender groups, teens' close friends' attitudes of indifference were associated with increased substance use and disapproval associated with reduced use, controlling for age, income, family structure, and adolescents' own attitudes of risk of substance use. Specifically, the effect of race on close friends' attitudes impacting cigarette and marijuana use was stronger for females compared to males. Female marijuana and cigarette use was more influenced by close friends' attitudes than males, and whites were more influenced by their close friends than Hispanics and blacks. It appears that white females were more susceptible to close friends' attitudes on cigarette use as compared to white males and youth of other races. These peer effects varied also by substance, with marijuana being the most peer-influenced substance and alcohol the least. Substance use prevention programming may need to emphasize different interventions with different subgroups (e.g., more peer based for white females, more family based for African–American and Hispanic teens). Regarding specific substances, for example, targeting peer influence on cigarette and marijuana use for white adolescents would make sense but would be less salient to Hispanics for cigarette, marijuana, and alcohol use.

In summary, it appears that both peer selection and influence effect substance use in adolescents and that particular substances moderate these differences. However, as the next section illustrates, recent genetically informed research is mounting evidence that the active selection process based on heritable characteristics for individuals is an important "child effect" dimension that needs to be a part of comprehensive peer network prevention research.

Peers, Genes, and Environments

Recent behavioral genetic research approaches provide powerful illumination on the mechanisms of influence and selection. Relative to this chapter, the peer group deviance (PGD) construct has been shown to be substantially heritable. In fact, as children gain agency in the development and management of their social lives, genetic influence on PGD appears to increase substantially and monotonically (Kendler, 2011; Kendler & Baker, 2007; Latendresse et al., 2011). For example, in a twin study that examined the moderating role of PGD on alcohol use for males at genetic risk for externalizing disorders, early adolescence appeared as a critical age. For 12–14 year olds, with low levels of PGD, the genetic risk for externalizing disorders had almost no impact on alcohol use; at high levels of PGD alcohol use

was strongly related to level of genetic risk (Kendler, 2001). These findings support earlier research that identified causal pathways for PGD via assortative friendship (Rose, 2002) and social selection or the "shopping model" (Dishion, Patterson, & Griesler, 1994; Patterson et al., 2000). This research points to interpretation that youth are actively seeking out others who share and reinforce their values and activities, with their genomes assisting in the creation of their social environments or reaching out "beyond their skin" to influence their social lives (Kendler, 2011). Thus, children's social lives are dependent to some extent on genetic factors, with studies indicating the role of genetics in peer processes such as friends' characteristics, peer group popularity, and peer victimization (Brendgen & Boivin, 2009). This comprehensive approach toward studying peer networks assists researchers in taking steps to clarify genetic and environmental influences. Because genetics are not destiny, the preventive implications for this research are important. For example, intervening in early adolescence (prior to age 15) may significantly attenuate the effect of elevated genetic risk for negative outcomes such as alcohol consumption. In all, this work is in its early stages and has much promise to inform peer network research design, analyses, interpretation, and policy implications.

Race/Ethnicity

In general, the limited findings addressing this issue of differences of peer effects by ethnicity have been mixed. For example, middle school-age children from Costa Rica nominated their parents as the primary provider of support, compared to similar aged US children who nominated their best friend as the most important source of support (DeRosier & Kupersmidt, 1991). Some research has found little to no differences in the role of peers on risk behavior as a function of ethnicity (Flannery, Vazsonyi, & Rowe, 1996; Urberg et al., 1997), while other research has found unique differences. For example, Giordano, Cernkovich, and DeMaris (1993) found that African–American adolescents reported less perceived peer pressure compared to European–Americans. This finding fits within a broader understanding of African–American youth receiving greater support from their extended and "fictive kin" family members and thus may be less influenced by their peers, compared to European–American youth (Taylor, Casten, & Flickinger, 1993). Peer influence has been studied with Mexican–American adolescents indicating that resistance to peer influence decreased as length of residence in the USA increased, possibly due to less emphasis on family as the primary source of influence (Bamaca & Umana-Taylor, 2006). In a study of Hispanic college students, friend and family support increased psychological well-being, but only friend support was predictive of lower psychological distress (Rodriguez, Mira, Myers, Morris, & Cardoza, 2003). Recent research has attempted to examine the role of ethnicity on peer association and has shown that positive indirect peer association may not function similarly among African–American, Hispanic, and European adolescents. Through the promotion of positive behaviors and protection against

risky or negative behaviors, the effects appear to be particularly relevant for European–American adolescents, when compared with African–American and Hispanic teens (Padilla-Walker & Bean, 2009). These differences were found only with indirect peer association, as no differences were found in relations between direct peer pressure and teen behavior by ethnicity. In all, ethnicity and cultural values are complex, important personal characteristics when studying and intervening with peer networks.

Sex Differences

As noted above, extensive research has shown the sex-based segregation of peer networks through childhood with more integration appearing in late middle and into high school for many adolescents (Brown, 1990; Brown & Klute, 2003; Rose & Rudolph, 2006; Vespo, Kerns, & O'Connor, 1996). The initial segregation has been labeled as the two-cultures theory, implying that boys and girls are essentially experiencing distinct socialization (Maccoby, 1998). Accordingly, girls were found to play in smaller groups that focused on intimate relationships, with boys playing in larger groups with more emphasis on sports or group activities (Neal, 2010). Network researchers have found mixed results with regard to size of networks, with no differences as found in network size between the sexes (Ennett & Bauman, 1996; Gest et al., 2007). In general, the network size difference appears to be moderated by age, with differences dissipating with age. The two-cultures theory also proposed differences in the density (the degree that a child's network peers know each other) of boys and girls networks, with girls having more dense networks compared to boys (Maccoby, 1998). Again, social network researchers have not found support for this distinction (Ennett & Bauman, 1996; Lee, Howes, & Chamberlain, 2007; Neal, 2010).

While the size and content of peer networks seems to progress in a fairly predictable pattern by sex differences, less is known about the mechanisms of adolescents' social networks that account for the influence of gender on the varying health outcomes such as substance use. Toward this end, recent research has begun to utilize a social–relational interpretative model in explaining differing substance use vulnerabilities between males and females. Socially informed substance abuse studies have explained substance use uptake for females through disturbances in social relationships that may influence self-medication (Wills, Sandy, & Yaeger, 2002; Zahn-Waxler, Crick, Shirtcliff, & Woods, 2006). This line of research maintains that as females are more likely to derive psychologically relevant information about themselves and others through interpersonal relationships, they are more vulnerable when they encounter interpersonal distress and therefore often experience increased disturbance when their relational ties are threatened, particularly with friends (Crick & Zahn-Waxler, 2003; Cross & Madson, 1997; Geary, 1998; Leadbeater, Blatt, & Quinlan, 1995; Maccoby, 1998). Related research analyzed whether the likelihood of urban adolescents' substance use involvement

was dependent on their social network quality (level of risk and protection) and whether that was moderated by gender and age. Results show young females' (13–16) substance use was strongly associated with their social networks compared to older females (17–20) and compared to young and older males (Mason et al., 2010). These findings highlight the varying affects of gender and age on adolescent social network protection and associated risk for outcomes such as substance use. In all, these findings point to the appropriateness of gender-sensitive social network research.

Contextual Approaches

In general, network researchers have routinely ignored context factors in explaining network structure, composition, and effects (Doreian & Conti, 2010). By studying social networks as a spatial entity or in a decontextualized manner, researchers have severely narrowed their lenses in examining social networks. For example, social network research on adolescent substance abuse has been largely approached through single settings to acquire complete sociometric network data, such as within school systems. These studies have contributed to our understanding of network structure and size and the association with substance use, yet this approach limits the capability to capture the unique contextual variations that make up an individual's socio-spatial life. An adolescent's social network may or may not consist of exclusively school-based peers, and the school setting is typically only one of many activity spaces that an adolescent frequents. By focusing only on one setting, researchers limit their ability to examine the socio-spatial variations within youths' lives. Research has demonstrated the interactive effects of the peer group and the setting for understanding outcomes. To address this issue, the construct activity space can provide theoretical and methodological guidance for addressing the socio-spatial dimension of youths' lives. Activity space can be defined as all the locations that an individual has direct contact with as a result of his or her daily activities (Miller, 1991). Activity spaces are the manifestation of individuals' spatial lives, serving as an index representing routine locations and all the accompanying psychological, social, and health-related experiences of these places (Golledge & Stimson, 1997; Sherman, Spencer, Preisser, Gesler, & Arcury, 2005). Recent research with urban youth demonstrates that the type of locations in which youth spend their time is varied and geographically dispersed and is not delimited by traditional geographical boundaries such as census tract, home neighborhood, block group, or political ward (Mason & Mennis, 2010). It is due to this unique spatial behavior of youth that traditional geographic boundaries or sample demarcations such as schools are not effective in capturing individuals' unique socio-spatial signatures and associated health outcomes. Recently, a new approach with peer networks has emerged that incorporates place as an essential ingredient when studying social networks.

Place-based social network research has found interactive linkages between place, peer networks, and health outcomes. For example, Mason and colleagues have found that substance use is dependent on the location and quality of peer social networks. In studying place-based social networks quality (levels of risk for and protection against substance use within a geographic context), they found that riskier networks at risky locations (poverty, crime, liquor stores, etc.) were related to more substance use, while protective networks at risky locations were related to less substance use (Mason et al., 2010). Substance use was increased by the presence of neighborhood concentrated disadvantage and interaction with substance-using peers at these risky places. Developmental interactive factors were also examined, with substance use among older adolescents more strongly related to social and geographic contexts than for younger adolescents, where the age range 14–16 appeared to be a critical transitional period for the emergence of social–contextual effects on substance use (Mennis & Mason, 2012).

This research asserts that social influences are not aspatial, but rather are embedded within place, and in fact play an important role in creating adolescents' experiences of place through the social interactions that occur at particular locations. Studies have established that individual perceptions of the characteristics of neighborhoods have been found to be a robust predictor of behavior such as substance use and mental health outcomes (Ellaway, Macintyre, & Kearns, 2002; Kawachi & Berkman, 2003; Lambert, Ialongo, Boyd, & Cooley, 2005; Latkin & Curry, 2003). Indeed, an adolescent's perception of a place is also dependent upon the people with whom the adolescent interacts and their attendant's attitudes, values, and behaviors at various locations (Mason, 2010).

By examining the influence of social networks on the interpretation of activity space, place-based network researchers have illuminated an interactive dynamic among peer networks, place, and health outcomes.

Location, as in real estate, is critical in understanding peer behaviors. Not only location, but the level of supervision, the type of activity, and the time of day are all important variables in predicting negative outcomes (Newman, Fox, Flynn, & Christenson, 2000). Youth who spend more time with peers in unstructured, unsupervised settings engage in more problem behaviors compared to youth who spend less (Osgood & Anderson, 2004). This important finding has recently been shown to be much more complicated. The type of structured activity (sports versus nonathletic activities) appears to interact with gender and produce differing effects. For example, in a large study of urban youth, delinquency was enhanced among boys who participated in sports compared to boys who participated in nonathletic activities, and not compared to boys who did not participate in any organized activities (Gardner et al., 2009). These variations among subgroups add to the complexity of understanding peers within a contextual framework.

Future Directions

Based upon the robust literature on peer networks, there is a reason to assume that this line of research will continue to develop and play an important role within prevention science. Like most areas within prevention science, peer network research is intricately linked to advances in statistics and technology. Leaders in the field of social network science are advancing more sophisticated analytical modeling techniques that expand the possible questions peer network researchers can address. Likewise, the simultaneous development of fast and powerful hardware and software platforms can leverage years of social science knowledge about networks and can enable researchers to model, test, and simulate network dynamics and evolutions in a more natural manner.

Economic learning models seek to understand network behaviors such as observation of others (social learning) in order to model social multiplier effects in the diffusion of innovations (Hogset & Barrett, 2010; Rogers, 1995). This approach has applications for prevention scientists interested in type-II (translational) research that seeks to move empirically supported interventions to scale. Individuals actively evaluate how interested others within their network are in innovative activities. Utilizing an agent-based, rational model has possible applications for researchers introducing behavioral innovations, activities, or the like that have preventive qualities targeting specific populations.

The explosive growth of technology related to communication and social networking and the implications for peer network research are underexploited. Interpersonal relations embedded within a technological space (e.g., Facebook) are shaped by the technology and are also shaping the fast-moving, adaptive technological changes. Examining online relational ties, multiplexed relationships, diffuse boundaries of space and time, and influence are critical areas that have yet to be fully leveraged by prevention scientists. Peer network research targeting the "online lives" of youth has enormous scientific potential to deepen the understanding of virtual and actual relationships, identities, and communities and how these relations shape attitudes and behaviors (Lewis, Kaufman, Gonzalez, Wimmer, & Nicholas Christakis, 2008; Watts, 2007). Adding a digital dimension to the domains when examining the relational practices of peers should now be integrated into network researchers' practice.

Prevention researchers concerned with peer networks must consider the fast-growing field of behavioral genetics. The debate over nature versus nurture is largely irrelevant, as the dynamic, interactive model has supplanted this older idea (Dick, 2011). Thus, we know that both genetic and environmental factors need to be accounted for by prevention scientists. Genes influence youth behavior at different levels during different age periods, which influence peer network formation. The microenvironmental context of families, schools, and neighborhoods provides varying levels of protection and risk for behavioral outcomes. All of these dynamic and interactive variables contribute to the complexity of this work but also shed more realism to the ongoing work to more accurately model peer

networks. Thus the need for prevention researchers to work broadly in an interdisciplinary capacity in order to comprehensively study the complex developmental movements associated with peer networks.

Conclusion

This review sought to underscore the importance of peer networks within prevention science, provide an overview of the history of this area, and examine current and future issues. The relational and developmental perspective necessary for the understanding of peer networks allows prevention scientist to logically theorize research problems within an ecological framework. Challenges for peer network research include expansion of research settings and thus populations and samples. Moving this area of research beyond single issue or disorder-focused studies such as substance abuse and HIV/AIDS is critical to continue to infuse the field with theory-driven lines of research. Integrating the strengths of multiple approaches (e.g., sociometric, egocentric, and place-based) with current developmental psychopathological and genetic research will expand and deepen peer network findings and preventive applications. For example, Bierman and Power (2009) recommend integrating emotional regulation and executive functioning research into peer relations studies in order to stimulate new social developmental applications. This type of peer relations research is clarifying the psychological and behavioral mechanisms of change and adaptation that occur within peer networks (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Greenberg, 2006). Networks are fluid and dynamic and exist within, between, and among multiple levels that are typically targeted in prevention science. The range of potential intervention points (full sociometric networks, personal networks, family, group, community, and individuals) to alter or modify peer networks toward more healthful relations provides prevention researchers with diverse backgrounds to engage in peer network research.

The particular issues addressed within the study of peer networks are vast and continue to grow. While network research is a complex area of study with multiple approaches and levels of knowledge required, at a basic level it is the study of how relationships among peers influence each other. This relational centerpiece should continue to be the focus for those interested in youth development. The need to further develop statistical and technological methods is particularly salient for peer network research due to the complexity of much of this research. Advances are also needed in observational and qualitative approaches to further ground and validate models and estimates with underrepresented populations. Development of contextual approaches that allow for the bringing to scale the micro-methodologies that have increased the understanding of place by network interactions needs to be advanced. The technological/online domain and its accompanying novel devices and applications have perhaps the greatest untapped potential in this growing area of study. Finally, the testing, adapting, and developing of peer network-related

theories that assist in explaining the complex pathways toward risk and prevention need to be aggressively promoted by prevention science and supported by funding entities. This theoretical advancement will provide a continued scientific scaffolding to advance the field. As our connectedness is related to our health and well-being, the study of peer networks continues to be an important and meaningful area of investigation for prevention science.

References

Bamaca, M. Y., & Umana-Taylor, A. J. (2006). Testing a model of resistance to peer pressure among Mexican-origin adolescents. *Journal of Youth and Adolescence, 35*, 631–645.

Bandura, A. (1996). Social cognitive theory of human development. In T. Husen & T. N. Postlethwait (Eds.), *International encyclopedia of education* (2nd ed., pp. 5513–5518). Oxford: Pergamon Press.

Barnes, J. A. (1954). Class and committees in a Norwegian island parish. *Human Relations, 7*, 39–58.

Barnes, J. A. (1969). Networks and political processes. In J. C. Mitchell (Ed.), *Social Networks in Urban Situations* (pp. 51–76). Manchester: Manchester University Press.

Barry, C. M., & Wentzel, K. R. (2006). Friend influence on prosocial behavior: The role of motivational factors and friendship characteristics. *Developmental Psychology, 42*, 153–163.

Batagelj, V., & Mrvar, A. (1998). Pajek: A program for large network analysis. *Connections, 21* (2), 47–57.

Bauman, K. E., & Ennett, S. T. (1996). On the importance of peer influence for adolescent drug use: Commonly neglected considerations. *Addiction, 91*(2), 185–198.

Bavelas, A. (1948). A mathematical model for group structure. *Human Organizations, 7*, 16–30.

Berndt, T. J. (2002). Friendship quality and social development. *Current Directions in Psychological Science, 11*(1), 7–10.

Berndt, T. J., Hawkins, J. A., & Jiao, Z. (1999). Influences of friends and friendships on adjustment to junior high school. *Merrill-Palmer Quarterly, 45*, 13–41.

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI Program. *Development and Psychopathology, 36*(8), 1175–1188.

Bierman, K. L., & Power, C. J. (2009). Social skills training to improve peer relations. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 603–621). New York, NY: Guilford Press.

Birch, S. H., & Ladd, G. W. (1996). Interpersonal relationships in the school environment and children's early school adjustment: The role of teachers and peers. In J. Juvonen & K. R. Wentzel (Eds.), *Social motivation: Understanding children's school adjustment* (pp. 199–225). New York, NY: Cambridge University Press.

Bolger, K. E., Patterson, C. J., & Kupersmidt, J. B. (1998). Peer relationships and self-esteem among children who have been maltreated. *Child Development, 69*, 1171–1197.

Borgatti, S. P. (1988). A comment on Doreian's regular equivalence in symmetric structures. *Social Networks, 10*, 265–271.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (1992). Ucinet IV: Network analysis software. *Connections, 15*, 12–15.

Brendgen, M., & Boivin, M. (2009). Genetic factors in children's peer relations. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 455–472). New York, NY: Guilford Press.

Bronfenbrenner, U. (1943). A constant frame of reference for sociometric research. *Sociometry, 6*, 363–397.

Brown, B. B. (1990). Peer groups and peer culture. In S. S. Feldman & G. R. Elliott (Eds.), *At the threshold* (pp. 171–196). Cambridge, MA: Harvard University Press.

Brown, B. B., & Klute, C. (2003). Friends, cliques, and crowds. In G. R. Adams & M. D. Berzonsky (Eds.), *Blackwell handbook of adolescence* (pp. 330–348). Malden, MA: Blackwell.

Byrne, D. (1971). The social structure of trust. *Social Networks, 20*, 265–289.

Cartwright, D., & Harary, F. (1956). Structural balance: A generalization of Heider's theory. *Psychological Review, 63*, 277–292.

Cattell, V. (2001). Poor people, poor places, and poor health: The mediating role of social networks and social capital. *Social Science & Medicine, 52*, 1501–1516.

Cho, H., Hallfors, D. D., & Sanchez, V. (2005). Evaluation of a high school peer group intervention for at-risk youth. *Journal of Abnormal Child Psychology, 33*(3), 363–374.

Christakis, N., & Fowler, J. (2009). *Connected*. New York, NY: Little, Brown and Company.

Cillessen, A. H. N. (2009). Sociometric methods. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 82–99). New York, NY: Guilford Press.

Cottrell, J. (2007). *Social networks in youth and adolescents*. London: Routledge.

Crick, N. R., & Zahn-Waxler, C. (2003). The development of psychopathology in females and males: Current progress and future challenges. *Development and Psychopathology, 15*, 719–742.

Cross, S. E., & Madson, L. (1997). Models of the self: Self construals and gender. *Psychological Bulletin, 122*, 5–37.

DeRosier, M. F., & Kupersmidt, J. B. (1991). Costa Rican children's perceptions of their social network. *Developmental Psychology, 27*, 656–662.

Dick, D. (2011). The dynamic nature of genetic and environmental influences of alcohol use and dependence. In K. Kendler, S. Jaffee, & D. Romer (Eds.), *The dynamic genome and mental health* (pp. 322–3339). Oxford: Oxford Press.

Dishion, T. J., & Andrews, D. W. (1995). Preventing escalation in problem behaviors with high-risk young adolescents: Immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology, 63*, 538–548.

Dishion, T. J., Dodge, K. A., & Lansford, J. E. (2006). Findings and recommendations: A blueprint in minimize deviant peer influence in youth interventions and programs. In K. A. Dodge, T. J. Dishion, & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions* (pp. 366–394). New York, NY: Guildford Press.

Dishion, T. J., Eddy, J. M., Haas, E., Li, F., & Spracklen, K. (1997). Friendships and violent behavior during adolescence. *Social Development, 6*, 207–223.

Dishion, T. J., Patterson, G. R., & Griesler, P. C. (1994). Peer adaptations in the development of antisocial behavior: A confluence model. In L. R. Huesmann (Ed.), *Aggressive behavior: Current perspectives* (pp. 61–95). New York, NY: Springer.

Dishion, T. J., Patterson, G. R., Stoolmiller, M., & Skinner, M. S. (1991). Family, school, and behavioral antecedents to early adolescent involvement with antisocial peers. *Developmental Psychology, 27*, 172–180.

Dishion, T. J., & Piehler, T. F. (2009). Deviant by design: Peer contagion in development, interventions, and schools. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 589–602). New York, NY: Guilford Press.

Dishion, T. J., Spracklen, K., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescent friendships. *Behavior Therapy, 27*, 373–390.

Doreian, P., & Conti, N. (2010). Social context, spatial structure and social network structure. *Social Networks*. doi:10.1016/j.socnet.2010.09.002.

Ellaway, A., Macintyre, S., & Kearns, A. (2002). Perceptions of place and health in socially contrasting neighbourhoods. *Urban Studies, 38*, 2299–2316.

Ennett, S. T., & Bauman, K. E. (1996). Adolescent social networks: School, demographic, and longitudinal considerations. *Journal of Adolescent Research, 11*, 194–215.

Epstein, J., & Botvin, G. (2002). The moderating role of risk-taking tendency and refusal assertiveness on social influences in alcohol use among inner-city adolescents. *Journal of Studies on Alcohol, 63*, 456–459.

Faris, R., & Ennett, S. (2012). Adolescent aggression: The role of peer group status motives, peer aggression, and group characteristics. *Social Networks, 34*(4), 371–378.

Festinger, L. (1954). A theory of social comparison process. *Human Relations, 7*, 1117–1140.

Fishbein, M., & Ajzen, I. (1981). Attitudes and voting behavior: An application of the theory of reasoned action. In G. M. Stephenson & J. M. Davis (Eds.), *Progress in applied social psychology* (Vol. I, pp. 253–313). London: Wiley.

Fisher, J. D., & Miscovich, S. J. (1990). Social influence and AIDS-preventing behaviors. In J. Edwards, R. Tindale, L. Heath, & E. Posavac (Eds.), *Applying social influence processes in preventing social problems*. New York, NY: Plenum Press.

Flannery, D. J., Vazsonyi, A. T., & Rowe, D. C. (1996). Caucasian and Hispanic early adolescent substance use: parenting, personality, and school adjustment. *Journal of Early Adolescence, 16*, 71–89.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*, 35–41.

Friedkin, N. E. (1998). *A structural theory of social influence*. Cambridge: Cambridge University Press.

Gardner, M., Roth, J., & Brooks-Gunn, J. (2009). Sports participation and juvenile delinquency: The role of the peer context among adolescent boys and girls with varied histories of problem behavior. *Developmental Psychology, 45*(2), 341–353.

Geary, D. C. (1998). *Male, female: The evolution of human sex differences*. Washington, DC: American Psychological Association.

Gest, S. D., Davidson, A. J., Rulison, K. L., Moody, J., & Welsh, J. A. (2007). Features of groups and status hierarchies in girls' and boys' early adolescent peer networks. In P. Rodkin & L. Hanish (Eds.), *New directions for child and adolescent development, Special Issue: Social network analysis and children's peer relationships*. San Francisco: Jossey-Bass.

Gifford-Smith, M., & Brownell, C. (2003). Childhood peer relationships: Social acceptance, friendships, and peer networks. *Journal of School Psychology, 41*, 235–284.

Giordano, P. C., Cernkovich, S. A., & DeMaris, A. (1993). The family and peer relations of Black adolescents. *Journal of Marriage and Family, 55*, 277–287.

Glaser, B., Shelton, K., & van den Bree, M. (2010). The moderating role of close friends in relationship between conduct problems and adolescent substance use. *Journal of Adolescent Health, 47*, 35–42.

Golledge, R., & Stimson, R. J. (1997). *Spatial behavior: A geographic perspective*. New York, NY: Guilford.

Greenberg, M. T. (2006). Promoting resilience in children and youth: Preventive interventions and their interface with neuroscience. *Annals of the New York Academy of Sciences, 1094*, 139–150.

Hartup, W. W. (2009). Critical issues and theoretical viewpoints. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 3–19). New York, NY: Guilford Press.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112*, 64–105.

Heaney, C. A., & Isreal, B. A. (1997). Social networks and social support. In K. Glanz, M. L. Frances, & R. K. Barbara (Eds.), *Health behavior and health education* (pp. 179–205). San Francisco, CA: Jossey-Bass Publishers.

Henrich, C. C., Kuperminc, G. P., Sack, A., Blatt, S. J., & Leadbeater, B. J. (2000). Characteristics and homogeneity of early adolescent friendship groups: A comparison of male and female cliques and non-clique members. *Applied Developmental Science, 4*, 15–26.

Hogset, H., & Barrett, C. (2010). Social learning, social influence, and projection bias: A caution on inferences based on proxy reporting of peer behavior. *Economic Development and Cultural Change, 58*(3), 563–589.

House, J., Umberson, D. J., & Landis, K. (1988). Structures and processes of social support. *Annual Review of Sociology, 14*, 293–318.

Hussong, A. M. (2002). Differentiating peer contexts and risk for adolescent substance use. *Journal of Youth and Adolescence, 31*, 207–220.

Katz, L., & Powell, J. H. (1955). Measurement of the tendency toward reciprocation of choice. *Sociometry, 18*, 659–665.

Kawachi, L. B., & Berkman, L. F. (2003). *Neighborhoods and health*. Oxford: Oxford Press.

Kendler, K. (2001). Twin studies of psychiatric illness: An update. *Archives of General Psychiatry, 163*, 433–440.

Kendler, K. (2011). Peer group deviance, conduct disorder, and alcohol intake. In K. Kendler, S. Jaffee, & D. Romer (Eds.), *The dynamic genome and mental health* (pp. 365–387). Oxford, UK: Oxford University Press.

Kendler, K., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine, 15*, 1–15.

Kiesner, J., Cadinu, M., Poulin, F., & Bucci, M. (2002). Group identification in early adolescence: Its relation with peer adjustment and its moderator effect on peer influence. *Child Development, 73*(1), 196–208.

Kirke, D. M. (2006). *Teenagers and substance use: Social networks and peer influence*. London: Palgrave Macmillan.

Lambert, S., Ialongo, N., Boyd, R., & Cooley, M. (2005). Risk factors for community violence exposure in adolescence. *American Journal of Community Psychology, 36*, 29–48.

Lansford, J. E., Capanna, C., Dodge, K. A., Capara, G. V., Bates, J. E., Pettit, G. S., et al. (2007). Peer social preference and depressive symptoms of children in Italy and the United States. *International Journal of Behavioral Development, 31*, 274–283.

Latendresse, S. J., Bates, J., Goodnight, J. A., Dodge, K. A., Lansford, J. E., & Pettit, G. S. (2011). Differential susceptibility to discrete patterns of adolescent externalizing behavior: Examining the interplay between CHRM2 and peer group deviance. *Child Development, 82*(6), 1797–1814.

Latkin, C. A., & Curry, A. D. (2003). Stressful neighborhoods and depression: A prospective study of the impact of neighborhood disorder. *Journal of Health and Social Behavior, 44*(1), 34–44.

Latkin, C. A., Mandell, W., Oziemkowska, M., Celentano, D., Vlahov, D., Ensminger, M., et al. (1995). Using social network analysis to study patterns of drug use among urban drug users at high risk for HIV/AIDS. *Drug and Alcohol Dependence, 38*, 1–9.

Lazarfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. In M. Berger, T. Abel, & C. H. Page (Eds.), *Freedom and control in modern society* (pp. 18–66). Princeton, NJ: Van Nostrand.

Leadbeater, B. J., Blatt, S. J., & Quinlan, D. M. (1995). Gender-linked vulnerabilities to depressive symptoms, stress, and problem behaviors in adolescents. *Journal of Research on Adolescence, 5*, 1–29.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *Journal of Abnormal and Social Psychology, 46*, 38–50.

Lee, L., Howes, C., & Chamberlain, B. (2007). Ethnic heterogeneity of social networks and cross-ethnic friendships of elementary school boys and girls. *Merrill-Palmer Quarterly, 53*, 325–346.

Leonard, K., Kearns, J., & Mudar, P. (2000). Peer networks among heavy, regular, and infrequent drinkers prior to marriage. *Journal of Studies on Alcohol, 61*, 669–673.

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Nicholas Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks, 30*(4), 330–342.

Maassen, G. H., van der Linden, J. L., Goosens, F. A., & Bokhorst, J. (2000). A ratings-based approach to two-dimensional sociometric status determination. In A. H. N. Cillessen & W. M. Bukowski (Eds.), *Recent advances in the measurement of acceptance and rejection in the peer system. New Directions for Child and Adolescent Development* (4th ed., pp. 1–101). New York, NY: Wiley.

Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together*. Cambridge, MA: Harvard University Press.

Mandell, W., Latkin, C. A., Oziemkowska, M., Vlahov, D., & Celentano, D. (1994). Changes in HIV risk behaviors among counseled injecting drug users. *Journal of Drug Issues, 24*, 555–567.

Marsden, P. (1990). Network data and measurement. *Annual Review of Sociology, 16*, 435–463.

Mason, M. J. (2009). Social network characteristics of urban adolescents in brief substance abuse treatment. *Journal of Child and Adolescent Substance Abuse, 18*, 72–84.

Mason, M. (2010). Attributing activity space as risky and safe: The social dimension to the meaning of place for urban adolescents. *Health and Place, 16*, 926–933.

Mason, M., & Mennis, J. (2010). An exploratory study of the effects of neighborhood characteristics on adolescent substance use. *Addiction Research & Theory, 18*(1), 33–50.

Mason, M., Mennis, J., Linker, J., Bares, C., & Zaharikas, N. (In Press). Peer attitudes effects on adolescent substance use: The moderating role of race and gender. *Prevention Science*

Mason, M., Valente, T., Coatsworth, J. D., Mennis, J., Lawrence, F., & Zelenak, P. (2010). Place-based social network quality and correlates of substance use among urban. *Journal of Adolescence, 33*(3), 419–427.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*, 414–444.

Mead, G. H. (1934). *Mind, self, and society*. Chicago, IL: University of Chicago Press.

Mennis, J., & Mason, M. (2012). Social and geographic contexts of adolescent substance use: The moderating effects of age and gender social networks. *Social Networks, 34*(1), 150–157.

Miller, H. (1991). Modeling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems, 5*, 287–301.

Mitchell, J. C. (Ed.). (1969). *Social networks in urban situations*. Manchester: Manchester University Press.

Molloy, L., Gest, S., & Rulison, K. (2011). Peer influences on academic motivation: Exploring multiple methods of assessing youths' most "influential" peer relationships. *Journal of Early Adolescence, 31*(1), 13–40.

Moreno, J. L. (1934). *Who shall survive? A new approach to the problem of human interrelations*. Washington, DC: Nervous and Mental Disease Publishing Company.

Mounts, N. S., & Steinberg, L. (1995). An ecological analysis of peer influence on adolescent grade point average and drug use. *Developmental Psychology, 31*, 915–922.

Neal, J. (2010). Hanging out: Features of urban children's peer social networks. *Journal of Social and Personal Relationships, 27*(7), 982–1000.

Newman, S. A., Fox, J. A., Flynn, E. A., & Christenson, W. (2000). *The prime time for juvenile crime or youth enrichment and achievement*. Washington, DC: Fight Crime Invest in Kids.

Oetting, E. R., & Donnermeyer, J. F. (1998). Primary socialization theory: The etiology of drug use and deviance. *Substance Use and Misuse, 33*(4), 995–1026.

Osgood, D. W., & Anderson, A. L. (2004). Unstructured socializing and rates of delinquency. *Criminology, 42*, 519–549.

Padilla-Walker, L., & Bean, R. (2009). Negative and positive peer influence: Relations to positive and negative behaviors for African American, European American, and Hispanic adolescents. *Journal of Adolescence, 32*, 323–337.

Parker, J. G., Rubin, K. H., Erath, S. A., Wojslawowicz, J. C., & Buskirk, A. A. (2006). Peer relationships, child development, and adjustment: A developmental psychopathology

perspective. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (pp. 419–493). Hoboken, NJ: John Wiley & Sons.

Patterson, G. R., Dishion, T. J., & Yoerger, K. (2000). Interpersonal dynamics within adolescent friendship: Dyadic mutality and deviant talk and patterns of antisocial behavior. *Prevention Science, 1*, 3–13.

Pearson, M., Steglich, C., & Snijders, T. (2006). Homophily and assimilation among sport-active adolescent substance users. *Connections, 27*(1), 47–63.

Pedersen, S., Vitaro, F., Barker, E. D., & Borge, A. (2007). The timing of middle childhood peer rejection and friendships: Linking early behaviors to adolescent adjustment. *Child Development, 78*, 1037–1051.

Piehler, T. F., & Dishion, T. J. (2007). Interpersonal dynamics within adolescent friendship: Dyadic mutuality and deviant talk and patterns of antisocial behavior. *Child Development, 78*(5), 1611–1624.

Rodriguez, N., Mira, C. B., Myers, H. F., Morris, J. K., & Cardoza, D. (2003). Family or friends: Who plays a greater supportive role for Latino college students? *Cultural Diversity and Ethnic Minority Psychology, 9*, 236–250.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York: Free Press.

Rose, R. J. (2002). How do adolescents select their friends? A behavior-genetic perspective. In L. Pulkkinen & A. Caspi (Eds.), *Paths to successful development: Personality in the life course* (pp. 106–128). Cambridge: Cambridge University Press.

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin, 132*, 98–131.

Schulenberg, J., Maggs, J. L., Dielman, T. E., Leech, S. L., Kloska, D. D., Shope, J. T., et al. (1999). On peer influences to get drunk: A panel study of young adolescents. *Merrill-Palmer Quarterly, 45*, 108–142.

Sherman, J. E., Spencer, J., Preisser, J. S., Gesler, W. M., & Arcury, T. A. (2005). A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics, 4*, 24. doi:10.1186/1476-072X-4-24.

Snijders, T. A. B., Steglich, C. E. G., Schweinberger, M., & Huisman, M. (2007) *Manual for SIENA version 3.1*. University of Groningen: ICS/Department of Sociology; University of Oxford: Department of Statistics.

Stattin, H., Gustafson, S. B., & Magnusson, D. (1989). Peer influences on adolescent drinking: A social transition perspective. *Journal of Early Adolescence, 9*, 227–246.

Steglich, C., Sniijders, T., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology, 40*(1), 329–393.

Sutherland, E. H., & Cressey, D. R. (1978). *Criminology*. Philadelphia, PA: J.B. Lippincott Company.

Taylor, R. D., Casten, R., & Flickinger, S. M. (1993). Influence of kinship social support on the parenting experiences and psychosocial adjustment of African-American adolescents. *Developmental Psychology, 29*, 382–388.

Terry, R., & Coie, J. D. (1991). A comparison of methods for defining sociometric status among children. *Developmental Psychology, 27*, 867–881.

Trice, H. M., & Sonnfnstuhr, W. J. (1988). Drinking behavior and risk factors related to the work place: Implications for research and preventions. *Journal of Applied Behavioral Science, 24*, 327–346.

Ueno, K. (2005). The effects of friendship networks on adolescent depressive symptoms. *Social Science Research, 34*(3), 484–510.

Urberg, K. A., Degirmencioglu, S. M., & Pilgrim, C. (1997). Close friend and group influence on adolescent cigarette smoking and alcohol use. *Developmental Psychology, 33*, 834–844.

Valente, T. (2003). Social networks' influence on adolescent substance use: An introduction. *Connections, 25*(2), 11–16.

Valente, T., Fujimoto, K., Chou, C., & Spruijt-Metz, D. (2009). Adolescent affiliations and adiposity: A social network analysis of friendships and obesity. *Journal of Adolescent Health, 45*, 202–204.

Valente, T., Unger, J., & Johnson, A. (2005). Do popular students smoke? The association between popularity and smoking among middle school students. *Journal of Adolescent Health, 37*, 323–329.

van Zalk, M., Kerr, M., Branje, S., Stattin, H., & Meeus, W. (2010). It takes three: Selection, influence, and de-selection processes of depression in adolescent friendship networks. *Developmental Psychology, 46*, 927–938.

Vespo, J. E., Kerns, K. A., & O'Connor, E. M. (1996). Organization in preschool classrooms: Social sets as social units. *Journal of Social and Personal Relationships, 13*, 41–55.

Vitaro, F., Boivin, M., & Bukowski, W. M. (2009). The role of friendship in child and adolescent psychosocial development. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 568–588). New York, NY: Guilford Press.

Wasserman, S. (1977). Random directed graph distributions and the triad census in social networks. *Journal of Mathematical Sociology, 5*, 61–86.

Watts, D. J. (2007). A twenty-first century science. *Nature, 445*, 489.

Wejnert, C. (2010). Social network analysis with respondent-driven sampling data: A study of racial integration on campus. *Social Networks, 32*, 112–124.

Wellman, B., & Berkowitz, S. D. (Eds.). (1988). *Social structures: A network approach*. Cambridge: Cambridge University Press.

Wentzel, K. R., Barry, C. M., & Cadwell, K. A. (2004). Friendships in middle school: Influences on motivation and school adjustment. *Journal of Educational Psychology, 96*, 195–203.

Wentzel, K. R., Filisetti, L., & Looney, L. (2007). Adolescent prosocial behavior: The role of self-processes and contextual cues. *Child Development, 78*, 895–910.

Wills, T. A., & Cleary, S. D. (1999). Peer and adolescent substance use among 6th-9th graders: Latent growth analyses of influence versus selection mechanisms. *Health Psychology, 19*(5), 453–463.

Wills, T. A., Sandy, J. M., & Yaeger, A. M. (2002). Stress and smoking in adolescence: A test of directional hypotheses. *Health Psychology, 21*(2), 122–130.

Zahn-Waxler, C., Crick, N., Shirtcliff, E., & Woods, K. (2006). The origins and development of psychopathology in females and males. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology. Theory & method* (Vol. 1, pp. 76–138). Hoboken, NJ: Wiley & Sons.

Chapter 9
Risk and Resilience Processes in Single-Mother Families: An Interactionist Perspective

Zoe E. Taylor and Rand D. Conger

Introduction

Risk and Resilience Processes in Single-Mother Families: An Interactionist Perspective

The developmental consequences of economic hardship and poverty for family functioning and child and adolescent adjustment continue to be of concern to developmentalists and policy makers. Economic changes in the United States during the past two decades, such as increasing income inequality, have renewed interest in how social position and economic resources affect families and the development of children (Conger & Donnellan, 2007). Research shows that more socially and economically disadvantaged adults and children are at higher risk for physical, emotional, and behavioral problems (Bradley & Corwyn, 2002; Conger, Conger, & Martin, 2010; McLoyd, 1998; Schoon, Sacker, & Bartley, 2003). Also important, it is increasingly evident that the family acts as a conduit for socioeconomic influences on the development of children and adolescents (e.g., Repetti, Taylor, & Seeman, 2002). Two dominant perspectives have been proposed to explain the nature of the relationships among contextual stressors, family functioning, and child development. The first, *social causation*, proposes that variations in environmental conditions, such as poverty or social support, lead to differences in social, emotional, cognitive, and physical functioning. In contrast the *social selection* perspective argues that differences in individual traits or characteristics account for the associations among adverse life events, family functioning, and life course development (Conger & Donnellan, 2007).

Z.E. Taylor (✉)
Department of Human Development and Family Studies, Purdue University, West Lafayette, IN, USA
e-mail: zetaylor@purdue.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_9,

Social Causation Perspectives

Two major theories are consistent with the social causation perspective. The Family Stress Model (FSM) describes how economic disadvantage exacerbates family stresses resulting in adverse family functioning, negative parenting, and poor child adjustment. The Family Investment Model (FIM) focuses on the ways in which parents invest financial, social, and human capital to promote children's physical, emotional, cognitive, and social well-being.

According to the FSM, economic pressure creates hardship-related emotions, behaviors, and conflicts for parents, which in turn negatively influence parenting and child outcomes (Conger et al., 2010; Conger & Donnellan, 2007). These processes negatively influence family functioning and individual adjustment primarily through the economic pressures they generate, such as the inability to pay bills and having to cut back on necessary expenses, such as health insurance. According to the FSM, these types of economic stressors negatively impact parenting quality by decreasing parental warmth, involvement, and support, and increasing depressive symptoms and marital conflict. Economic stress indirectly affects child adjustment as parents who are distracted by their own problems show less affection and support to their children and are more harsh, irritable, and inconsistent with their parenting practices. These problems increase children's risk of internalizing (i.e., symptoms of depression and anxiety) and externalizing (aggressive, deviant, or antisocial) behaviors. The processes described in the FSM are especially pertinent for single-mother families whom have a high risk for economic adversity. However, the FSM has largely been tested with two-parent families (for exceptions see Mistry, Vandewater, Huston, & McLoyd, 2002; Taylor, Larsen-Rife, Conger, Widaman, & Cutrona, 2010; Taylor, Widaman, Robins, Jochem, Early, & Conger, 2012). Research is also needed that considers factors that might buffer families from the adverse processes described in the FSM (Conger et al., 2010).

The second theory relating to a social causation perspective is the FIM, which Conger and Donnellan (2007) proposed as an extension of the investment model in economics (Coleman & James, 1988). According to the FIM, families with higher socioeconomic status (SES) have greater access to financial, human (e.g., education and skills), and social (connections to and the status and power of individuals in one's social network) capital. The model proposes that these family resources are associated with the successful development of children and adolescents, as families with greater economic resources are able to make significant investments in the development of their children, whereas more disadvantaged families must invest in more immediate family needs (Bradley & Corwyn, 2002; Conger & Donnellan, 2007). Thus, the FIM proposes that parents with greater resources are likely to invest their economic, educational, and occupational capital in ways that facilitate the positive development of their children from childhood into adulthood. The FSM and FIM are especially relevant for single-mother families as these families often face relatively disadvantaged circumstances and lower SES.

Social Selection Perspectives

The social selection approach to explaining the association between family stress processes and family functioning proposes that these associations result from individual differences (e.g., Rowe & Rodgers, 1997). According to this viewpoint, some individual characteristics facilitate the accumulation of social advantages (e.g., high SES), whereas others generate risk for disadvantages (e.g., conflict in social relationships). Behavioral genetic studies have shown that up to 50 % of the variance in personality characteristics results from genetic influences (Bouchard & McGue, 2003; Loehlin, Neiderhiser, & Reiss, 2003). Several investigators hypothesize that genetic differences can foster vulnerability to risk and adversity (e.g., Rowe & Rodgers, 1997). Rowe and Rodgers (1997) hypothesize that individual differences in heritable traits are both passed on to offspring and are reflected in an individual's parenting style. Importantly, selection effects acknowledge the influence of biological characteristics that parents bring to their families of procreation. Consistent with the assumptions of social selection, research reported during the past decade has demonstrated that individual traits and dispositions during childhood and adolescence predict to later status attainment (Conger et al., 2010). Evidence from longitudinal studies has revealed that early emerging individual differences in personality, aggressiveness, and cognitive ability predict economic and relationship outcomes in adulthood (e.g., Donnellan, Conger, McAdams, & Neppl, 2009; McLeod & Kaiser, 2004; Shiner, Masten, & Roberts, 2003). Thus, research examining social selection effects is also important for explaining developmental adjustment and functioning across the life span.

Single-Mother Families

Mother-headed families appear to be particularly vulnerable to the consequences of economic hardship. Research shows that individuals in single-mother families are susceptible to a variety of social and contextual risks ranging from economic hardship to poor psychological functioning (i.e., Avison, Ali, & Walters, 2007; Brown & Moran, 1997; Murry, Bynum, Brody, Willert, & Stephens, 2001). Poverty and high levels of stress are particularly troublesome for single mothers and result in an above-average risk for emotional distress and disruptions in effective parenting (i.e., Avison, 1999; Fergusson, Boden, & Horwood, 2007). In turn, children from single-parent families are at increased risk of poor outcomes largely due to exposure to more adverse environments resulting from poverty, parental internalizing symptoms, and ineffective parenting methods (Amato & Keith, 1991; Fergusson et al., 2007; McLoyd, 1998). Thus, the social causation and social selection perspectives may be particularly useful in better understanding the adjustment of single-mother families. However, there remain important limitations in the research focused on single-parent families.

First, although research has documented the risks and vulnerabilities associated with single parenthood relatively few studies have examined how single mothers often are resilient or overcome these vulnerabilities (for exceptions see Cain & Combs-Orme, 2005; Hanson, 1986; Murry & Brody, 1999; Murry et al., 2001; Taylor et al., 2010; Taylor et al., 2012). Resiliency is a process of positive adaption in the context of significant threats to an individuals' development or functioning (Masten & Wright, 2009). Studies of resilience provide evidence that positive outcomes are possible for individuals experiencing even extreme hardships (Luthar, 2006). Examining the processes that contribute to resiliency in vulnerable families is crucial for identifying pathways that disrupt maladaptive outcomes and promote competent development during hard times or difficult transitions (Cicchetti & Toth, 2009; Conger, Belsky, & Capaldi, 2009; Patterson, 2002).

Second, studies have typically compared single-mother families to two-parent families and have not examined within-group differences that likely exist among mother-headed families. Although earlier research has identified many of the general vulnerabilities that single mothers face, much less is known about the pathways or mechanisms that contribute to poor family functioning, how these negative pathways can be disrupted, or whether risks are equally distributed among subgroups of single-mother families. Single mothers of various racial backgrounds may also have important risk and resilience differences that largely remain to be determined (Garcia Coll, Lamberty, Jenkins, McAdoo, & Crnic, 1996; Gibson-Davis & Gassman-Pines, 2010; Luthar, 2006). For example, African Americans and Hispanics may be more likely than other ethnic groups to utilize extended kin networks for support during difficult times (e.g., Oropesa & Landale, 2004).

Last, developmental research is moving toward an integrated, multidisciplinary perspective that addresses developmental processes across time and considers diversity in developmental pathways (Cicchetti & Toth, 2009). Bioecological models suggest that children's development occurs in nested multilevel contexts, from biological to cultural, with risk factors occurring across levels (Bronfenbrenner & Ceci, 1994). The mechanisms involved in associations between risk and resilience in single-mother families most likely involve multiple social, biological, and behavioral factors (Repetti et al., 2002). Despite these priorities in the field of family studies, little research has used a multidisciplinary approach to improve understanding of functioning in single-mother families, especially across the life course.

An Interactionist Model of Single-Mother Families

A model that may be particularly useful for guiding and advancing research on single mothers, as well as for addressing prior limitations, is the interactionist perspective proposed by Conger and Donnellan (2007). An interactionist perspective integrates social causation (environmental) and social selection (biological or individual difference) perspectives by proposing that the relationship between

social conditions and individual adjustment involves a dynamic interplay between these two processes (Avison, 1999; Conger & Donnellan, 2007). The strength of an interactionist approach is that it helps overcome the limitations of either the social causation or selection model. That is, social selection perspectives tend to minimize the impact of environmental influences on individual well-being, whereas social causation perspectives often underestimate the influence of individual differences on social and emotional development (Conger & Donnellan, 2007). The interactionist model acknowledges that both social selection and social causation processes influence individual development across the life course. Although social causation and social selection have often been viewed as competing theories in regard to the origins of adversity and psychopathology, recent studies using the interactionist model have demonstrated the existence of a dynamic interplay between processes of social selection and social causation. Consistent with the theory, this research has shown that individual dispositions affect SES and family stress processes that, in turn, predict the development of individual characteristics (Conger et al., 2010; Martin et al., 2010; Schofield et al., 2011; Trentacosta et al., 2010). For example, Martin et al. (2010) found that both social selection and social causation processes accounted for associations among SES, individual development, and family functioning.

An interactionist perspective is potentially useful for increasing understanding of mother-headed families as it provides a theoretical framework for consideration of both environmental and individual difference factors that potentially contribute to resilient family processes. Specifically, an interactionist perspective encourages consideration of both the strengths and the vulnerabilities that may be present in single-mother families, acknowledges that risk and resilience processes can be associated with both environmental as well as dispositional factors, and simultaneously examines these processes across time.

Consistent with these ideas, we propose an interactionist model for advancing research on single-mother families that includes proposed causal pathways that reflect both a social causation and social selection perspective (Fig. 9.1). Our model begins with the mother's developmental history, such as biological and psychological characteristics, that is likely to contribute to the current circumstances and quality of family relationships and in turn child adjustment. The model we propose reflects various processes and pathways (e.g., current resources and stressors) that mediate the current circumstances and quality of family relationships and eventually child adjustment in the current family. In addition to mediating pathways, we expect moderating pathways that we will discuss later. Importantly, this model allows examination of both strengths and vulnerabilities present in single-mother families. Next we review research relevant to single-mother families integrating findings from both a social selection and a social causation framework. We consider each construct in the model and its associations with the other components organized around the model outlined in Fig. 9.1.

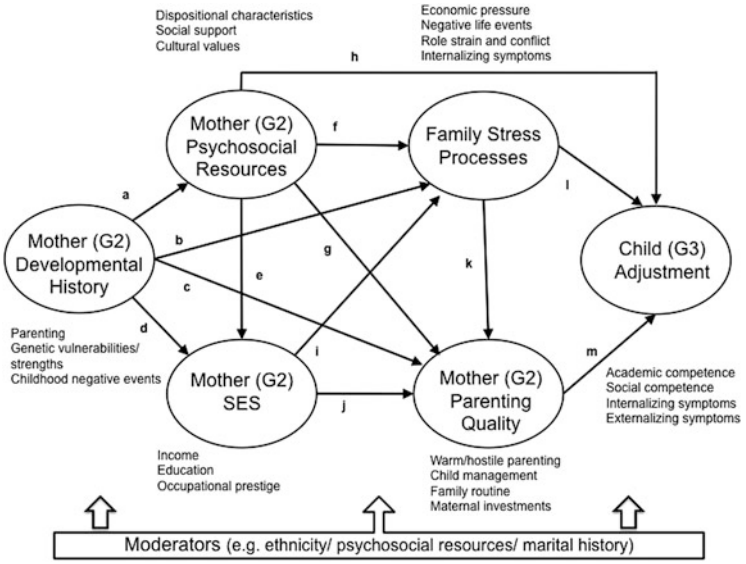

Fig. 9.1 An interactionist model of risk and resilience in single-mother families

Mother's Developmental History

The interactionist model (Fig. 9.1, paths a–d) begins with the mother's developmental history, such as biological and psychological characteristics likely to contribute to later functioning in single-mother families. Taking into account an individual's early developmental history highlights the role that cumulative risk or advantages have over the life course (Sroufe, Carlson, Levy, & Egeland, 1999). At times, the higher risks prevalent in single-mother families may be the result of an underlying series of social, educational, economic, and familial disadvantages that begin early in life (Fergusson et al., 2007). For example, important dimensions of developmental history include parenting, genetic strengths and vulnerabilities, and early childhood events. Our first path (Fig. 9.1, path a) predicts that developmental history is connected to the psychosocial resources available to mothers in adulthood. Psychosocial resources can include physiological, psychological, or environmental resources that are available to an individual such as individual characteristics, family strengths, cultural factors, and community supports. Parenting quality and behaviors in one's family of procreation are both associated with short- and long-term developmental outcomes across a variety of domains such as psychological and physiological functioning and social and cognitive competence (Repetti et al., 2002; Sroufe et al., 1999). Early adverse experiences can have cascading negative influences on mental and physical health and behavioral trajectories across the life span by creating vulnerability to future stress or by exacerbating biologically based vulnerabilities (Curtis & Cicchetti, 2003; Repetti et al., 2002).

Research has rarely addressed whether developmental history predicts psychosocial resources in single mothers. However, consistent with this hypothesis, Taylor et al. (2010) reported that childhood adversity negatively predicted dispositional optimism in single African American mothers.

We also predict that developmental history is associated with current family stress processes and investment in children (paths b and c). Research indicates that some single mothers are at risk for maladjustment, such as internalizing symptoms or economic pressure, because they experienced high levels of stressors and negative events earlier in life (Cairney, Boyle, Offord, & Racine, 2003; Davies, Avison, & McAlpine, 1997; Fergusson et al., 2007; Taylor et al., 2010). Studies show that the number of childhood adversities experienced by single mothers increases their risk for early onset of depression as well as increases the risk for subsequent depressive episodes (Avison et al., 2007; Cairney et al., 2003; Davies et al., 1997). Single mothers more often report that their own parents had serious problems with alcohol, drugs, and depression (Avison et al., 2007; Cairney et al., 2003) and have higher rates of childhood abuse than married mothers (Lipman, MacMillan, & Boyle, 2001).

Genetic strengths and vulnerabilities may also predict later family stress and parenting behaviors. Studies have identified genes that protect maltreated children from psychopathology later in their life (Caspi et al., 2002) as well as a gene that moderates the effects of stressful life events on depression (Caspi et al., 2003). The social selection perspective is consistent with findings suggesting that some individuals have characteristics that make them more at risk for experiencing marital discord, divorce, and single parenthood (Avison, 1999; Lahey et al., 1988; Simons, Beaman, Conger, & Chao, 1993). For example, Lahey et al. (1988) found that divorced mothers had higher levels of an antisocial behavior trait in comparison to married mothers. Studies have also shown that antisocial behavior has an indirect association with parenting quality and psychological distress, economic strain, and negative life events (Simons et al., 1993) and that mothers with higher levels of antisocial behavior report more marital transitions than those with lower levels (Capaldi & Patterson, 1991). This research demonstrates important links between an individual's functioning and neurobiological systems that have clear implications for research addressing psychopathology as well as resilience in vulnerable families.

Our final path in regard to developmental history predicts an association between mother's developmental history and current SES (path d). Developmental risks, such as poverty, are often passed on to successive generations (Conger et al., 2009). For example, being born into a disadvantaged family increases the probability of cascading negative life events and decreased academic achievement—processes that are often reflected in later adult achievements and SES (Schoon et al., 2003). Conversely, growing up in an economically well-to-do family confers human capital in terms of greater education and income later in life. Single mothers with such early advantages should have greater personal and material resources for coping with single parenthood.

Psychosocial Resources and Family Adjustment

The next pathways in our model are concerned with the effects of psychosocial resources on mothers' current circumstances (paths e–g). Enduring social or psychological resources plays an important function in mediating or moderating the pernicious influence of current life stressors (Avison, 1999; Kessler & Essex, 1982). Psychological and social resources such as attitudes, skills, or assets impact overall family functioning by influencing mental health and positive parenting behaviors and buffering individuals from adversity. Important psychosocial resources identified in Fig. 9.1 can include positive dispositional characteristics, environmental factors such as social support, and cultural or spiritual values. These resources can benefit single mothers by alleviating some of the family stress processes commonly occurring in single-mother families such as current SES (path e); family stressors such as economic pressure, negative life events, and depression (path f); and parenting quality (path g).

Dispositional Resources

Certain traits or personal characteristics, such as dispositional optimism, self-efficacy, and self-esteem, have been positively linked to adjustment in single-mother families. Dispositional optimism is a relatively stable, general tendency of individuals to expect positive outcomes (Carver, Scheier, & Segerstrom, 2010). Research shows that individuals high in optimism have better psychological and physical adjustment to negative life events and typically report less distress across a broad range of stressful situations, have more positive social networks, and enjoy better physical health (Carver et al., 2010; Nes & Segerstrom, 2006; Taylor & Stanton, 2007). Other research has shown that maternal optimism predicted lower levels of maternal internalizing symptoms and higher levels of effective parenting for both African American and Mexican-origin single mothers (Taylor et al., 2012; Taylor et al., 2010). Studies have also found that optimism is associated with positive coping and parenting behaviors in African American single-mother families, which predict higher child cognitive competence, social competence, and psychological adjustment across time (i.e., Brody, Murry, Kim, & Brown, 2002; Kim & Brody, 2005; Taylor, Rodriguez, Seaton, & Dominguez, 2004; Taylor et al., 2010).

Self-efficacy and self-esteem are other dispositional characteristics that are associated with better adjustment in single-mother families. Feelings of self-efficacy can determine an individual's coping strategies, how much effort she will expend, and for how long (Bandura, 1977). Studies with African American families show that self-efficacy is an important resource for single mothers especially in regard to predicting feelings of well-being (i.e., Brody, Flor, & Gibson, 1999; Jackson, 2000; Jackson & Schemes, 2005; Olson, Ceballo, & Park, 2002). Self-efficacy has also been linked to mothers' reports of positively handling

emotional stressors (Olson et al., 2002). Self-esteem also influences parenting behaviors and effectiveness. Brody and colleagues found that higher self-esteem was positively associated with family routines, mother–child relationship quality, and more effective parenting across time in African American single-mother families (i.e., Brody et al., 1999, 2002; Jackson, 2000; Jackson & Schemes, 2005; Olson et al., 2002). However, the high rates of stressors often experienced by single mothers can erode their self-esteem and reduce their feeling in control of their lives (Avison, 1999; Mandara, Johnston, Murray, & Varner, 2008).

Dispositional resources may also be positively associated with SES in single mothers. Consistent with this hypothesis, Schofield et al. (2011) found evidence that alpha personality (high levels of agreeableness and conscientiousness and a low level of neuroticism) in adolescence predicted higher levels of SES in adulthood. However, research on the effects of dispositional and personality characteristics on SES in single-mother families remain scarce.

Social Resources

In addition to individual characteristics, we propose that social resources will play an important role in how well single mothers function. Social support refers to one's social bonds, social integration, and primary group relations and is well recognized to be indispensible to personal functioning and psychological well-being throughout the life course, especially in regard to depression (Turner, 1999). Lack of social support is associated with single mothers' internalizing symptoms (Hilton & Kopera-Frye, 2006; Kotchick, Dorsey, & Heller, 2005; Murry et al., 2001). However, internalizing symptoms tend to decrease as perceived social support increases (Turner, 1999). Social support networks can provide important financial, emotional, and childcare support for single mothers. In particular, kinship support may be especially meaningful for single mothers in terms of affecting parenting and child and adolescent outcomes (i.e., Burchinal, Follmer, & Bryant, 1996; Taylor & Roberts, 1995). For example, social support from relatives is associated with increased feelings of well-being for African American single mothers and their children (Jackson, 2000; Taylor & Roberts, 1995). Single African American mothers who rely on grandmothers for child care report higher quality parenting practices, improved psychological adjustment, and higher self-efficacy, which in turn predicts positive child adjustment (Murry et al., 2001). Harknett (2006) additionally found that mothers with the most social network support worked more, had higher incomes, and received less in welfare benefits than mothers with the least support. However, studies examining kin social support and well-being of single-mother families have largely been conducted with African American families, and research with other ethnic groups is needed.

Research is inconclusive about whether single mothers benefit from living with kin. A study found that African American children do better with a grandmother in the household compared with households with only a single adult (Simons, Chen, Simons, Brody, & Cutrona, 2006). However, it is also noteworthy that some

research shows that kin social support, in particular support from grandmothers, is associated with family conflict, maternal depression, and parenting stress in African American families (for a review see Murry et al., 2001). This may be due to poor interpersonal relationships between the mother and her daughter, in that family involvement can encompass both positive support and negative interference (Murry et al., 2001). More research is needed to determine under what circumstances social support from kin is either detrimental or helpful and whether these relationships are moderated by ethnicity.

Cultural and Spiritual Resources

Researchers have begun to examine cultural and spiritual beliefs and traditions as protective factors that promote positive outcomes for individuals and families (Cardoso & Thompson, 2010; Cauce & Domenech Rodriguez, 2002). We propose that cultural values will be associated with functioning in single-mother families, although little research has examined this hypothesis. For example, consistent with literature that cites the importance of cultural values such as familism, Mexican-origin parents are more likely to rely on family members for instrumental and emotional support in comparison with other ethnic groups (Cauce & Domenech Rodriguez, 2002; Schaffer & Wagner, 1996; Wagner, 1987). In addition, Mexican-origin single mothers typically have larger social support networks compared with Caucasian single mothers (Schaffer & Wagner, 1996). Spirituality and religiosity are associated with greater resilience in families and can reinforce ethnic and cultural pride (Cardoso & Thompson, 2010). Unfortunately, examination of whether spiritual values serve as a resilience factor has largely involved two-parent families. However, Greeff and Fillis (2009) did find that religiosity predicted higher psychological functioning in single mothers. Few researchers have examined how cultural or spiritual beliefs are associated with parenting behaviors in single-mother families. One exception is a study that found that maternal religiosity predicted mother–child relationship quality and maternal school involvement in a sample of African American single mothers (Brody & Flor, 1997).

Child Adjustment

The final path in our model proposes that mothers' psychosocial resources can directly predict child adjustment (path h), although there are most likely indirect effects via family stress and parenting quality. A social resource that may be particularly meaningful for child adjustment in single-mother families is whether noncustodial fathers are involved with their children. Creating a harmonious co-parenting relationship following a divorce or breakup has important implications for single mothers, noncustodial fathers, and especially their children. Single mothers with joint custody report lower levels of parenting stress and better co-parenting relationships, which in turn is associated with positive adjustment in

children (Arditti & Madden-Derdich, 1997). In contrast, less supportive and hostile co-parenting relationships increase child emotional and behavioral problems (Arditti & Bickley, 1996; Hetherington & Stanley-Hagan, 1999). Prior marital status appears to differentiate father involvement, with divorced fathers maintaining higher levels of contact than never-married fathers (McKenry, McKelvey, Leigh, & Wark, 1996).

However, father involvement may not uniformly benefit single-mother families. Studies report mixed findings regarding whether nonresident father involvement buffers adolescent problem behaviors across ethnic groups. Pan and Farrell (2006) found that nonresident father involvement decreased delinquency for African American adolescents but that nonresident father involvement in Asian American and Hispanic single-mother families was associated with increased adolescent problem behavior. However, Thomas, Farrell, and Barnes (1996) found that African American male adolescents in single-mother families had fewer problem behaviors when nonresident fathers were not involved. Thomas et al. (1996) also reported that Caucasian male adolescents in single-mother families without the support of a nonresident father had higher rates of problem behavior than African American adolescents. Overall, more research is needed to differentiate between when father involvement is beneficial, and when it is not, in single-mother families. In addition, although a growing body of research suggests that cooperative co-parenting has significant benefits for single-mother families, this is an ideal that isn't typically attained. Research addressing how to improve the quality of the co-parenting relationship, reduce conflict, and foster fathers' positive involvement is vital (Cowan, Cowan, Pruett, & Pruett, 2007).

SES and Family Functioning

From SES to Family Stress Processes

Our model also addresses the pathways between SES and family adjustment. SES typically reflects three interrelated indicators: education, income, and occupational status. Our model proposes that SES will affect family stress processes (Fig. 9.1, path i). SES has important implications for individual and families' mental health, physical well-being, and achievements across the life course (Bradley & Corwyn, 2002). A strong association exists between family structure and socioeconomic disadvantage. Single mothers' family income and assets are typically the lowest of all household groups (Duncan, Brooks-Gunn, & Klebanov, 1994; Harknett, 2006; Lino, 1995). Single mothers also are more likely to have less education and be younger in age than married mothers (Cairney et al., 2003). Furthermore, although poverty has increased for all racial groups during the past decade, ethnic minorities remain disproportionately poor, and a larger percentage of mother-headed Hispanic and African American households live below the poverty line compared with

Caucasian (non-Hispanic) mother-headed households (U.S. Bureau of the Census, 2009). Minority children are significantly more likely to be born into an unmarried family or to a never-married mother without a partner (Gibson-Davis & Gassman-Pines, 2010). Mexican immigrant single mothers also face significant socioeconomic challenges such as lower levels of education and employment and much higher rates of poverty compared with US-born single mothers (Padilla, Radey, Hummer, & Kim, 2006).

However, socioeconomic resources are associated with lower levels of internalizing symptoms in single mothers (Brody et al., 2002). Studies with African American single mothers found that working combined with maternal education and higher income was negatively associated with financial strain (Gyamfi, Brooks-Gunn, & Jackson, 2001; Jackson, 2000). Educational advancement also improves the economic status and, in turn, psychological functioning of single-mother families (Zhan, 2007). In addition, financial resources can serve as a protective factor against adversity during a divorce. A study of African American families found that mothers with greater financial resources did not increase in depressive symptoms after a divorce, compared with mothers with fewer resources (Varner & Mandara, 2009).

The model also proposes that SES predicts mothers' investments in their children (path j). Family resources are associated with the successful development of children and adolescents; families with greater economic resources are able to make significant investments in the development of their children, whereas more disadvantaged families must invest in more immediate family needs (Conger & Donnellan, 2007). The standard of living for mothers and their children decreases significantly after a divorce, and a large percentage of single mothers do not receive child support from their former spouses (Duncan et al., 1994). However, single mothers' educational attainment is typically associated with more supportive and involved parenting (Brody & Flor, 1997; Jackson, 2000).

Family Stress Processes

Our model is particularly concerned with the effects of family stressors on parenting and child adjustment (Fig. 9.1, paths k and l). The capacity for supportive, sensitive, and involved parenting is often compromised in emotionally distressed, disadvantaged families. In turn, disruptions of positive parenting are a significant risk factor for child maladjustment (McLoyd, 1990). As previously discussed, single parents are at higher risk for inadequate and ineffective parenting behaviors largely due to higher levels of life stressors, which predict poorer psychological functioning and disrupt family routines, relationships, and parenting behaviors (e.g., Brody & Flor, 1997; Kim & Brody, 2005; Mistry et al., 2002).

A considerable stressor for single-mother families is economic pressure, as previously discussed in the FSM (Conger et al., 2010). Studies have demonstrated that single mothers typically experience high levels of economic pressure that results in poor family functioning such as poor parenting behaviors, maternal

internalizing symptoms, and poor child adjustment (Mistry et al., 2002; Taylor et al., 2010; Taylor et al., 2012). Taylor et al. (2010) reported that mother's childhood adversities, current economic pressure, and internalizing problems were associated with lower levels of maternal warmth and child management and with lower child school competence across time. Similarly, Mistry et al. (2002) found that economic pressure indirectly affected parenting by adversely impacting mother's psychological functioning. Distressed parents were observed to be less affectionate and reported feeling less effective and capable in disciplinary interactions with their children.

Role strain, or the dual demands of work and childcare responsibilities, also has a significant impact on stress processes in single-mother families. Studies show that work demands increase work–family conflict, negatively impact family routines, and increase maternal symptoms of depression (Gyamfi et al., 2001; McLoyd, Toyokawa, & Kaplan, 2008). In particular, research shows that low-wage jobs often increase role strain without greatly reducing economic strain, resulting in continued poor functioning (Zhan, 2007). However, higher educated mothers who are better off financially also report greater levels of parenting stress, perhaps as a result of work spilling over into their family life (Arditti & Bickley, 1996).

Taken together, researchers have demonstrated that the environmental conditions experienced by single mothers, such as poverty and low SES, result in emotional strain and poor mental health, which in turn are significant risk factors for ineffective parenting and negative child adjustment. An important aspect of our model is that it highlights the need for research to examine what processes buffer or lessen the impact of family stressors prevalent in single-mother families, specifically the influence of psychosocial resources on family stress processes and parenting quality as discussed above.

From Parenting to Child Adjustment

The final path (Fig. 9.1, path m) proposes that the quality of parenting will affect child adjustment across a variety of developmental domains such as academic competence, internalizing symptoms, and externalizing behaviors. Maternal characteristics such as warmth, responsiveness, efficacy, and family routine have been shown to buffer children from adverse situations (for a review, see Luthar, 2006). Children with supportive parents have higher self-esteem, fewer behavioral problems, more effective coping skills, and better mental health and are physically healthier than children with non-supportive parents (Gardner & Cutrona, 2004). In contrast, studies consistently demonstrate that children with unsupportive and harsh parents have lower self-esteem, more psychological problems, higher behavior problems, and lower prosocial behavior (Cain & Combs-Orme, 2005; Cairney et al., 2003; Ceballo & McLoyd, 2002; McLoyd, 1990; Simons et al., 1993). Thus, research that examines factors that are associated with quality parenting and strong family relationships in spite of adverse circumstances is of utmost importance.

Moderating Pathways in the Interactionist Model

So far we have discussed research linked to compensatory resources or those resources that directly counteract adversity. Compensatory variables have a direct or main effect on outcomes of theoretical interest. For example, we expect that parental optimism will be positively related to supportive parenting, thus counteracting in part the negative effects of socioeconomic disadvantage. However, there are also potential protective factors that alter or attenuate the relationships between adversity and individual development or the family processes described above. Very few studies have tested moderating pathways related to adversity in single-mother families. Important moderating factors that deserve more attention regarding single-mother families include prior marital history, race and ethnicity, and psychosocial resources.

Prior Marital History as a Moderator

Specific groups of single parents may face unique circumstances. However, research on events and experiences that precede single parenthood and contribute to later marital circumstances is limited (Avison, 1999). Single mothers who are never married differ in demographics, income, and mental health when compared with divorced single mothers (Afifi, Cox, & Enns, 2006; Jayakody, Chatters, & Taylor, 1993). It is also likely that poor mental health and other risk factors are predicted by different mechanisms for different subgroups of single mothers such as those who are single mothers through divorce or those who never married. For example, never-married mothers may be at higher risk for experiencing depression as a result of poverty, whereas divorced mothers may be at risk due to loss of income and partnership. Consistent with this possibility, a study of ethnically diverse mothers found that separated and divorced mothers had higher levels of depression and generalized anxiety disorder compared with never-married mothers, whereas never-married mothers were generally more similar to married mothers in their mental health profiles (Afifi et al., 2006). Transitional periods may also be associated with differing levels of adversity and strain. Hetherington and Stanley-Hagan (1999) found that divorced mothers report significant emotional distress for the first few years after a divorce but that well-being and parenting quality improve as family roles and relationships restabilize. This finding suggests that there may be important differences between single mothers based on marital history that needs further examination and clarification in future research.

Race and Ethnicity as a Moderator

The associations between family functioning and stress most likely differ by race or ethnicity as a result of group differences in wealth and assets, cultural norms, and

informal kinship support (Gibson-Davis & Gassman-Pines, 2010). For example, in contrast to literature with mainly Caucasian families, studies have found that mothers' household structure did not affect parenting stress or practices in African American families (Cain & Combs-Orme, 2005; Gibson-Davis & Gassman-Pines, 2010). However, more research addressing these findings is needed. It is also important to note that what constitutes beneficial parenting behaviors may differ as a result of environmental context. Murry et al. (2001) found that parenting constructs differ in African American families due to contextual roles and factors such as discrimination or dangerous, risky neighborhoods. African Americans used a more no-nonsense style of parenting characterized by higher levels of warmth than a typical authoritarian parenting style but higher levels of control, monitoring, and vigilance. Differences in parenting styles for Latino families have also been found. Domenech Rodríguez, Donovick, and Crowley (2009) found that the dominant parenting style for first-generation Latino families was a "protective" style that was high in warmth, high in demandingness, but low in autonomy granting.

Cultural beliefs have also been found to moderate levels of social support given to single mothers by their relatives. Mexican-origin and African American mothers received less support from their relatives if they were divorced rather than never-married (Jayakody et al., 1993; Wagner, 1987). In contrast European-origin single mothers received less support from their relatives if they were never married rather than divorced (Wagner, 1987). Taylor, Hurley, and Riley (1986) additionally found that less acculturated Hispanic single mothers relied on family support more than more acculturated single mothers.

Psychosocial Resources as a Moderator

Most studies of single-mother families have examined mediating links between a positive construct, such as social support or personality, and family functioning. In contrast, little research has tested whether these resources attenuate, or moderate, negative pathways that result in poor functioning in single-parent families. An exception to this is work by Taylor and colleagues (2010, 2012) that has examined whether dispositional optimism attenuates the relationship between economic pressure and internalizing symptoms in African American and Mexican-origin single-parent families. They found that optimism had little effect on the relationship between economic pressure and internalizing symptoms when economic pressure was low. However, optimism had an increasingly positive effect as economic pressure increased. When economic pressure was high, mothers with relatively higher levels of optimism had lower levels of internalizing symptoms than those who were less optimistic. Other studies have shown that self-esteem buffered individuals from experiencing internalizing symptoms, as a result of adversity, whereas low self-esteem creates more vulnerability to adversity (Brown & Moran, 1997; Costello, Compton, Keeler, & Angold, 2003). For example, Brown and Moran (1997) found that an onset of depression was more likely to follow an

adverse event when women had poor self-esteem and lacked social support, both of which were commonly reported by single mothers.

Conclusions and Future Directions

Future Research and Theory

We have presented a new heuristic framework for guiding and advancing research on single-mother families. As described earlier, our theoretical model utilizes both social selection and social causation perspectives. In terms of social selection, the model utilizes variables addressing personal resources and characteristics such as positive temperament and optimism, as well as how these characteristics can be passed onto future generations, i.e., the path between family of origin characteristics to mother's psychological resources and from mother's psychological resources to child outcomes. In relation to social causation perspectives, the model addresses how family stress processes, both in the family of origin as well as currently, negatively affect family functioning in mother-headed families. Further research is needed to assess what psychological, social, and environmental factors assist in promoting positive functioning and resilience in mother-headed families. The model we have presented reflects this focus, by allowing for a more complete assessment of the risks of single motherhood, as well as allowing for a concurrent assessment of protective factors.

Equally important for future research is that differences among single-mother families need more attention. Despite circumstances that most likely vary considerably, single parents are often viewed as a homogenous group when their trajectories of emotional well-being most likely differ as a result of a variety of factors. Acknowledging single parents as a multifaceted group deserving to be studied in their own right allows for improved insight into both the strengths and weakness inherent in this type of family structure and would allow for more effective prevention and intervention efforts. For example, some single mothers may experience poor emotional functioning as a result of economic strain and stress, whereas for others, poor psychological functioning may be the result of adversity in childhood. Especially important for prevention will be the examination of the developmental trajectories leading to single parenthood and the ways in which developmental background interacts with current risks to increase vulnerability as well as resilience in these families.

Other differences may also be important. A marital transition such as divorce may also have a positive influence in reducing adversity if women are experiencing an abusive or unhappy marriage. Divorced women are less depressed, less anxious, feel more independent, and have fewer health problems than women in high-conflict non-divorced families (Hetherington, 1993). Research has also not differentiated between single mothers whose ex-partner remains involved in parenting

versus those who have no contact with their children's father. We also know very little about what differentiates fathers who remain involved with their children after a separation or divorce from fathers who do not. Finally, as discussed, single mothers are not homogenous in terms of cultural and racial differences. Research addressing race and ethnic differences in single-mother families in terms of both risk and resiliency factors is clearly needed. Minority single mothers may face even more unique stressors than have been described in this review. For example, single mothers who are immigrants may experience language difficulties, problems relating to acculturation, and social isolation. Together, this model provides a significantly more integrative assessment of functioning in single-mother families and allows for the complexities of single-mother families to be examined.

Implications for Prevention and Intervention

Our model suggests several points of entry for the development of more effective intervention programs working with increasing positive functioning in single-mother families. Also important to note is that although our model emphasizes functioning in single-mother families, the model has implications for vulnerable families or families in general. First, interventions that target reducing stressful experiences early in development are vital for lessening exposure to adversity in adulthood and preventing some of the cascading risks hypothesized in our model (Fig. 9.1). Research has highlighted how childhood adversities create a cascade of risk and vulnerability for individuals that last into adulthood and detrimentally influence their own families (Repetti et al., 2002; Taylor et al., 2010). It is important to recognize that some single mothers have had lives characterized by frequent adversities that have impaired positive functioning across their life span (Avison, 1999). These risks are often passed on to future generations. Thus, interventions that promote positive parenting behaviors and that foster resilience to adversity are crucial not just for individual adjustment but for the current development of children and, critically, for adjustment across their life span. It is increasingly evident from longitudinal studies that stressors and adversity (as well as resources and resilience) influence children's present environment, but also have long-term developmental consequences that last into adulthood and continue on in parenting the next generation.

In terms of current stressors, interventions that target economic circumstances, parenting behaviors, and individual dispositions and resources might be particularly beneficial for single mothers. Although most individuals experience stressful events at some point during their lifetime, what stands out for single mothers is they often experience persistent stressors as well as multiple simultaneous stressors (Avison, 1999). The magnitude of stressors affecting single mothers often results from being both the primary caregiver and the primary wage earner in their families, and these stressors are compounded by prior adverse events across the life span.

Most important for prevention research is garnering an understanding of the mechanisms by which these risks can be ameliorated.

Finally, as discussed, research has most often focused on negative aspects of functioning in single-mother families. Research examining the influence of psychosocial resources within single-mother families will be crucial for determining interventions that not only prevent negative outcomes but that encourage positive ones. An example of this is the work on dispositional optimism and functioning in single-mother families described earlier by Taylor et al. (2010, 2012). Although dispositional optimism is a relatively stable trait, research indicates that optimism is linked to effective coping strategies that can be fostered through psychosocial intervention (Carver et al., 2010; Peterson, 2000). Modifying coping strategies has the potential to help individuals manage stress and avoid compromising their mental health (Taylor & Stanton, 2007).

A More General View

Single mothers lead complex lives that are shaped by the combined influence of social causation and social selection across their life span. These challenges influence emotional health and parenting behaviors, as well as the developmental outcomes of their children. The most important next step for researchers interested in alleviating adjustment problems in single-mother families is to examine the processes linked to resilience and competence in spite of risk and vulnerabilities. Research that explores ways to reduce the impact of specific risks associated with single parenthood, that curbs negative chain reactions across the life span, and that promotes positive traits and resources could have important implications for positive family functioning and ameliorating risk which could then be utilized in implementing intervention programs. Research that moves beyond describing the risks inherent in single parenthood to examining the ways in which single mothers successfully cope with the challenges they often face would have important implications for policy, interventions, and most importantly for families themselves.

References

Afifi, T. O., Cox, B. J., & Enns, M. W. (2006). Mental health profiles among married, never-married, and separated/divorced mothers in a nationally representative sample. *Social Psychiatry and Psychiatric Epidemiology, 41*, 122–129.

Amato, P. R., & Keith, B. (1991). Parental divorce and the well-being of children: A meta-analysis. *Psychological Bulletin, 110*, 26–46.

Arditti, J., & Bickley, P. (1996). Father's involvement and mothers' parenting stress postdivorce. *Journal of Divorce & Remarriage, 26*, 1–23.

Arditti, J., & Madden-Derdich, D. (1997). Joint and sole custody mothers: Implications for research and practice. *Families in Society, 78*, 36–45.

Avison, W. R. (1999). Family structure and processes. In A. V. Horwitz & T. L. Scheid (Eds.), *A handbook for the study of mental health: Social contexts, theories, and systems* (pp. 1228–1240). New York: Cambridge University Press.

Avison, W. R., Ali, J., & Walters, D. (2007). Family structure, stress, and psychological distress: A demonstration of the impact of differential exposure. *Journal of Health and Social Behavior, 48*, 301–317.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215.

Bouchard, T. J., Jr., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology, 54*, 4–45.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*, 371–399.

Brody, G. H., & Flor, D. L. (1997). Maternal psychological functioning, family processes, and child adjustment in rural, single-parent, African-American families. *Developmental Psychology, 33*, 1000–1011.

Brody, G. H., Flor, D. L., & Gibson, N. M. (1999). Linking maternal efficacy beliefs, developmental goals, parenting practices, and child competence in rural single-parent African American families. *Child Development, 70*, 1197–1208.

Brody, G. H., Murry, V. M., Kim, S., & Brown, A. C. (2002). Longitudinal pathways to competence and psychological adjustment among African-American children living in rural single-parent households. *Child Development, 73*, 1505–1516.

Bronfenbrenner, U., & Ceci, S. I. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review, 101*, 568–586.

Brown, G. W., & Moran, P. M. (1997). Single mothers, poverty and depression. *Psychological Medicine, 27*, 21–33.

Burchinal, M. R., Follmer, A., & Bryant, D. M. (1996). The relations of maternal social support and family structure with maternal responsiveness and child outcomes among African-American families. *Developmental Psychology, 32*, 1073–1083.

Cain, D. S., & Combs-Orme, T. (2005). Family structure effects on parenting stress and practices in the African-American Family. *Journal of Sociology and Social Welfare, 32*, 19–40.

Cairney, J., Boyle, M., Offord, D., & Racine, Y. (2003). Stress, social support and depression in single and married mothers. *Social Psychiatry and Psychiatric Epidemiology, 38*, 442–449.

Capaldi, D. M., & Patterson, G. R. (1991). Relation of parental transitions to boys' adjustment problems: I. A linear hypothesis. II. Mothers at risk for transitions and unskilled parenting. *Developmental Psychology, 27*, 489–504.

Cardoso, J. B., & Thompson, S. (2010). Common themes of resilience among Latino immigrant families: A systematic review of the literature. *Families in Society, 91*, 257–265.

Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review, 30*, 879–889.

Caspi, A., McClay, J., Moffitt, T., Mill, J., Martin, J., Craig, I. W., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science, 297*, 851–854.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., et al. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science, 301*, 386–389.

Cauce, A. M., & Domenech Rodriguez, M. (2002). Latino families: Myths and realities. In J. Contreras (Ed.), *Latino youth and families facing the future*. New York: Basic Books.

Ceballo, R., & McLoyd, V. C. (2002). Social support and parenting in poor, dangerous neighborhoods. *Child Development, 73*, 1310–1321.

Cicchetti, D., & Toth, S. L. (2009). The past achievements and future promises of developmental psychopathology: The coming of age of a discipline. *Journal of Child Psychology and Psychiatry, 50*, 16–25.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*, S95–S120. Suppl.: Organizations and Institutions: Sociologicaland Economic Approaches to the Analysis of Social Structure.

Conger, R., Belsky, J., & Capaldi, D. M. (2009). The intergenerational transmission of parenting: Closing comments for the special section. *Developmental Psychology, 45*, 1276–1283.

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and Family, 72*, 685–704.

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annual Review of Psychology, 58*, 175–199.

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *JAMA, 290*, 2023–2029.

Cowan, C. P., Cowan, P. A., Pruett, M. K., & Pruett, K. (2007). An approach to preventing coparenting conflict and divorce in low-income families: Strengthening couple relationships and fostering fathers' involvement. *Family Process, 46*, 109–121.

Curtis, J. W., & Cicchetti, D. (2003). Moving research on resilience into the 21st century: Theoretical and methodological considerations in examining the biological contributors to resilience. *Development and Psychopathology, 15*, 773–810.

Davies, L., Avison, W. R., & McAlpine, D. D. (1997). Significant life experiences and depression among single and married mothers. *Journal of Marriage and the Family, 59*, 294–308.

Domenech Rodríguez, M. M., Donovick, M. R., & Crowley, S. L. (2009). Parenting styles in a cultural context: Observations of "protective parenting" in first-generation Latinos. *Family Process, 48*, 195–210.

Donnellan, M. B., Conger, K. J., McAdams, K. K., & Neppl, T. K. (2009). Personal characteristics and resilience to economic hardship and its consequences: Conceptual issues and empirical illustrations. *Journal of Personality, 77*, 1–31.

Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child Development, 65*, 296–318.

Fergusson, D. M., Boden, J. M., & Horwood, J. (2007). Exposure to single parenthood in childhood and later mental health, educational, economic, and criminal behavior outcomes. *Archives of General Psychiatry, 64*, 1089–1095.

Garcia Coll, C., Lamberty, G., Jenkins, R., McAdoo, H. P., & Crnic, K. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development, 67*, 1891–1914.

Gardner, K. A., & Cutrona, C. E. (2004). Social support communication in families. In A. L. Vangelisti (Ed.), *Handbook of family communication* (pp. 495–512). Mahwah, NJ: Lawrence Erlbaum.

Gibson-Davis, C. M., & Gassman-Pines, A. (2010). Early childhood family structure and mother-child interactions: Variation by race and ethnicity. *Developmental Psychology, 46*, 151–164.

Greeff, A. P., & Fillis, A. J. (2009). Resiliency in poor single-parent families. *Families in Society, 90*, 279–285.

Gyamfi, P., Brooks-Gunn, J., & Jackson, A. P. (2001). Associations between employment and financial and parental stress in low-income single black mothers. *Women and Health. Special Issue: Welfare, work, and well-being: Part I, 32*, 119–135.

Hanson, S. M. (1986). Healthy single parent families. *Family Relations, 35*, 125–132. Special Issue: The Single Parent Family.

Harknett, K. (2006). The relationship between private safety nets and economic outcomes among single mothers. *Journal of Marriage and Family, 68*, 172–191.

Hetherington, E. M. (1993). An overview of the Virginia Longitudinal Study of Divorce and Remarriage with a focus on the early adolescent. *Journal of Family Psychology, 7*, 39–56.

Hetherington, E. M., & Stanley-Hagan, M. (1999). The adjustment of children with divorced parents: A risk and resiliency perspective. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 129–140.

Hilton, J. M., & Kopera-Frye, K. (2006). Loss and depression in cohabiting and non-cohabiting custodial single parents. *The Family Journal, 14*, 28–40.

Jackson, A. P. (2000). Maternal self-efficacy and children's influence on stress and parenting among single black mothers in poverty. *Journal of Family Issues, 21*, 3–16.

Jackson, A. P., & Schemes, R. (2005). Single mothers' self-efficacy, parenting in the home environment, and children's development in a two-wave study. *Social Work Research, 29*, 7–20.

Jayakody, R., Chatters, L. M., & Taylor, R. J. (1993). Family support to single and married African American mothers: The provision of financial, emotional, and child care assistance. *Journal of Marriage and the Family, 55*, 261–276.

Kessler, R. C., & Essex, M. (1982). Marital status and depression: The importance of coping resources. *Social Forces, 61*, 484–507.

Kim, S., & Brody, G. (2005). Longitudinal pathways to psychological adjustment among Black youth living in single-parent households. *Journal of Family Psychology, 19*, 305–331.

Kotchick, B. A., Dorsey, S., & Heller, L. (2005). Predictors of parenting among African American single mothers: Personal and contextual factors. *Journal of Marriage and Family, 67*, 448–460.

Lahey, B. B., Hartdagen, S. E., Frick, P. J., Mc-Burnett, K., Connor, R., & Hynd, G. W. (1988). Conduct disorder: Parsing the confounded relation to parental divorce and antisocial personality. *Journal of Abnormal Psychology, 97*, 334–337.

Lino, M. (1995). The economics of single parenthood: Past research and future directions. *Marriage and Family Review. Special Issue: Single parent families: Diversity, myths and realities: I, 20*, 99–114.

Lipman, E. L., MacMillan, H. L., & Boyle, M. H. (2001). Childhood abuse and psychiatric disorders among single and married mothers. *American Journal of Psychiatry, 158*, 73–77.

Loehlin, J. C., Neiderhiser, J. M., & Reiss, D. (2003). The behavior genetics of personality and the NEAD study. *Journal of Research in Personality, 37*, 373–387.

Luthar, S. S. (2006). Resilience in development: A synthesis of research across five decades. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology, Vol 3: Risk, disorder, and adaptation* (2nd ed., pp. 739–795). Hoboken, NJ: Wiley.

Mandara, J., Johnston, J. S., Murray, C. B., & Varner, F. (2008). Marriage, money, and African American mothers' self-esteem. *Journal of Marriage and Family, 70*, 1188–1199.

Martin, M. J., Conger, R. D., Schofield, T. J., Dogan, S. J., Widaman, K. F., Donnellan, M. B., Neppl, T. K. (2010). Evaluation of the interactionist model of socioeconomic status and problem behavior: A developmental cascade across generations. *Development and Psychopathology. Special Issue: Developmental cascades: Part 1, 22,* 695–713.

Masten, A. S., & Wright, O. M. (2009). Resilience over the lifespan: Developmental perspectives on resistance, recovery, and transformation. In J. W. Reich (Ed.), *Handbook of adult resilience* (pp. 213–237). New York: Guilford Publications, Inc.

McKenry, P. C., McKelvey, M. W., Leigh, D., & Wark, L. (1996). Nonresidential father involvement: A comparison of divorced, Separated, never married, and remarried fathers. *Journal of Divorce and Remarriage, 25*, 1–13.

McLeod, J. D., & Kaiser, K. (2004). Childhood emotional and behavioral problems and educational attainment. *American Sociological Review, 69*, 636–658.

McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socio-emotional development. *Child Development, 61*, 311–346.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53*, 185–204.

McLoyd, V. C., Toyokawa, T., & Kaplan, R. (2008). Work demands, work-family conflict, and child adjustment in African American families: The mediating role of family routines. *Journal of Family Issues, 29*, 1247–1267.

Mistry, R. S., Vandewater, E. A., Huston, A. C., & McLoyd, V. C. (2002). Economic well-being and children's social adjustment: The role of family process in an ethnically diverse low-income sample. *Child Development, 73*, 935–951.

Murry, V. M., & Brody, G. H. (1999). Self-regulation and self-worth of Black children reared in economically stressed, rural, single mother-headed families: The contribution of risk and protective factors. *Journal of Family Issues, 20*, 458–484.

Murry, V. M., Bynum, M. S., Brody, G. H., Willert, A., & Stephens, D. (2001). African American single mothers and children in context: A review of studies on risk and resilience. *Clinical Child and Family Psychology Review, 4*, 133–155.

Nes, L. S., & Segerstrom, S. C. (2006). Dispositional optimism and coping: A meta-analytic review. *Personality and Social Psychology Review, 10*, 235–251.

Olson, S. L., Ceballo, R., & Park, C. (2002). Early problem behavior among children from low-income, mother-headed families: A multiple risk perspective. *Journal of Clinical Child and Adolescent Psychology, 31*, 419–430.

Oropesa, R. S., & Landale, N. S. (2004). The future of marriage and Hispanics. *Journal of Marriage and Family, 66*, 901–920.

Padilla, Y. C., Radey, M. D., Hummer, R. A., & Kim, E. (2006). The living conditions of US born children of Mexican immigrants in unmarried families. *Hispanic Journal of Behavioral Sciences, 28*, 331–349.

Pan, E., & Farrell, M. P. (2006). Ethnic differences in the effects of intergenerational relations on adolescent problem behavior in U.S. single mother families. *Journal of Family Issues, 27*, 1137–1158.

Patterson, J. M. (2002). Integrating family resilience and family stress theory. *Journal of Marriage and Family, 64*, 349–360.

Peterson, C. (2000). The future of optimism. *American Psychologist, 55*, 44–55.

Repetti, R. L., Taylor, S. E., & Seeman, T. E. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin, 128*, 330–366.

Rowe, D. C., & Rodgers, J. L. (1997). Poverty and behavior: Are environmental measures nature and nurture? *Developmental Review, 17*, 358–375.

Schaffer, D. M., & Wagner, R. M. (1996). Mexican-origin and Anglo single-mothers: The influence of ethnicity, generation, and socioeconomic status on social support networks. *Hispanic Journal of Behavioral Sciences, 18*, 74–86.

Schofield, T. J., Martin, M. J., Conger, R. D., Neppl, T. K., Donnellan, M. B., & Conger, K. J. (2011). Intergenerational transmission of adaptive functioning: A test of the interaction model of SES and human development. *Child Development, 82*, 33–47.

Schoon, I., Sacker, A., & Bartley, M. (2003). Socioeconomic adversity and psychosocial adjustment: A developmental-contextual perspective. *Social Science and Medicine, 57*, 1001–1015.

Shiner, R. L., Masten, A. S., & Roberts, J. M. (2003). Childhood personality foreshadows adult personality and life outcomes two decades later. *Journal of Personality, 71*, 1145–1170.

Simons, R. L., Beaman, J., Conger, R. D., & Chao, W. (1993). Stress, support, and antisocial behavior trait as determinants of emotional well-being and parenting practices among single mothers. *Journal of Marriage and the Family, 55*, 385–398.

Simons, L. G., Chen, Y., Simons, R., Brody, G., & Cutrona, C. (2006). Parenting practices and child adjustment in different types of households: A study of African American families. *Journal of Family Issues, 27*, 803–825.

Sroufe, L. A., Carlson, E. A., Levy, A. K., & Egeland, B. (1999). Implications of attachment theory for developmental psychopathology. *Development and Psychopathology, 11*, 1–13.

Taylor, V. L., Hurley, E. C., & Riley, M. T. (1986). The influence of acculturation upon the adjustment of preschool Mexican-American children of single-parent families. *Family Therapy, 13*, 249–256.

Taylor, Z. E., Larsen-Rife, D., Conger, R. D., Widaman, K. F., & Cutrona, C. E. (2010). Life stress, maternal optimism, and adolescent competence in single-mother, African-American families. *Journal of Family Psychology, 24*, 468–477.

Taylor, R. D., & Roberts, D. (1995). Kinship support and maternal and adolescent well-being in economically disadvantaged African-American families. *Child Development, 66*, 1585–1597.

Taylor, R. D., Rodriguez, A. U., Seaton, E. K., & Dominguez, A. (2004). Association of financial resources with parenting and adolescent adjustment in African-American families. *Journal of Adolescent Research, 19*, 267–283.

Taylor, S. E., & Stanton, A. L. (2007). Coping resources, coping processes, and mental health. *Annual Review of Clinical Psychology, 3*, 377–401.

Taylor, Z. E., Widaman, K. F., Jochem, R., Early, D., & Conger, R. D. (2012). Dispositional optimism: A psychological resource for Mexican-origin mothers experiencing economic stress. *Journal of Family Psychology, 26*(1), 133–139.

Thomas, G., Farrell, M. P., & Barnes, G. M. (1996). The effects of single mother families and nonresident fathers on delinquency and substance abuse in Black and White adolescents. *Journal of Marriage and the Family, 58*, 884–894.

Trentacosta, C. J., Neppl, T. K., Donnellan, M. B., Scaramella, L. V., Shaw, D. S., & Conger, R. D. (2010). Adolescent personality as a prospective predictor of parenting: An interactionist perspective. *Journal of Family Psychology, 24*, 721–730.

Turner, R. J. (1999). Social support and coping. In A. V. Horwitz & T. L. Scheid (Eds.), *A handbook for the study of mental health: Social contexts, theories, and systems* (pp. 1198–1210). New York: Cambridge University Press.

U.S. Bureau of the Census. (2009). *Historical income data: Current population survey tables*. Retrieved April 12, 2010, from http://www.census.gov/hhes/www/poverty/poverty08.html.

Varner, F., & Mandara, J. (2009). Marital transitions and changes in African American mothers' depressive symptoms: The buffering role of financial resources. *Journal of Family Psychology, 23*, 839–847.

Wagner, R. M. (1987). Changes in extended family relationships for Mexican-origin and Anglo single-mothers. *Journal of Divorce, 11*, 69–87. Special Issue: Minority and ethnic issues in the divorce process.

Zhan, M. (2007). Assets, human capital development, and economic mobility of single mothers. *Families in Society, 88*, 605–615.

Chapter 10
Environmental Influences: The Workplace and Mental Health—Models, Vulnerability Factors, and Interventions

Stephen Stansfeld, Sokratis Dinos, and Joel Raffel

Introduction

Much of adult life is spent in the workplace. Working provides many benefits in terms of financial resources, social status, boosting self-esteem, and providing a social network. However, the workplace can have negative as well as positive effects on mental health and behavioral disorders. High levels of demands at work without the compensation of control over work, an imbalance between effort invested and reward received, bullying and discrimination at work, and job insecurity have been related to a greater risk of common mental disorders. Conversely, social support from supervisors and colleagues can have positive effects on mental health.

Health effects in the workplace vary according to position in the life course, and there are a number of markers of potential vulnerability. Research has shown that some groups are at high risk of mental health problems, and the workplace can convey an additional risk. Risk markers include gender, ethnicity, sexuality, previous mental health problems, and learning disabilities. Those starting work and older people may also be vulnerable.

This chapter presents the relatively few studies that have assessed organizational workplace interventions tackling primary, secondary, and tertiary prevention of ill health. There are considerable limitations to current research on the workplace and health, and many areas need further development and more robust methodology. The identification of effective workplace-based interventions to improve the mental health of the economically active is at its infancy and requires considerable further research.

S. Stansfeld (✉)
Centre for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, Old Anatomy Building, London EC1M 6BQ, UK
e-mail: s.a.stansfeld@qmul.ac.uk

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_10,

The workplace is an important potential site for the prevention of ill health. In industrialized countries the majority of the population spends most of its adult life working. Thus, what happens in the workplace is influential for people's lives and well-being. Work provides many useful benefits: financial resources, the social status of an occupation, a structure to the day, boosts to self-esteem, and a social network with whom to interact. Not working or being unemployed, carries well-recognized risks to health including increased risk of depressive and anxiety disorders, demoralization, and suicide (Bartley, 1997; Fryers, Melzer, & Jenkins, 2003).

One characteristic of the workplace is that it often implies shared exposure to positive or negative influences for health. In this chapter, we critically examine the main models that have linked the work environment to mental health and illness, the association of specific occupations with mental health, and the ways in which mental ill health may manifest itself in the workplace. We then document research that has identified populations that may be particularly vulnerable to workplace stressors. We consider research on gender, sexuality, ethnicity, age, learning disabilities, and preexisting mental health, in this context. Finally, we examine interventions in the workplace, at both the individual and the organizational levels that have been carried out with a view to preventing mental ill health. We look at the strengths and limitations of different interventions strategies, as well as the limitations of the research that has evaluated them.

Psychosocial Work Environment

Several models have been devised to understand the hazards and benefits of the psychosocial work environment and its link to risks of physical and mental ill health. One of the most influential models has been Karasek's Job Strain Model, also called the Demand–Control Model (Karasek & Theorell, 1990) (Fig. 10.1). This model comprises two key dimensions: psychological work demands and decision latitude. Decision latitude is made up of control over work (decision authority), the variety of work, and opportunity for use of skills (skill discretion). According to the Job Strain Model, jobs can be classified into four types:

1. The most risky types of job for health are high-strain jobs typified by high psychological demands and low decision latitude.
2. Low-strain jobs with few psychological demands and high levels of decision latitude are predicted to have lower than average levels of psychological strain and lower risk of ill health because there are relatively few challenges, and decision latitude allows the worker to respond optimally to these few challenges.
3. Active jobs with high levels of demands and high levels of decision latitude can be rewarding but also may be associated with higher levels of mental ill health because of the relatively potent effect of high demands.

Fig. 10.1 Karasek's job strain model

	Decision Latitude	
	HIGH	LOW
HIGH	"Active"	"High strain"
LOW	"Low strain"	"Passive"

4. Passive jobs typified by low demands and low control can be demotivating and involve the atrophy of skills and abilities, but only average levels of psychological strain and health risk are expected.

A refinement of this model is the addition by Johnson, Hall, and Theorell (1989) of the further dimension of occupational social support. The revised model, so-called Demand–Control–Support or Iso-Strain Model, includes support received from supervisors and coworkers that may buffer the effect of high demands and low control. Within this model, decision latitude (or control over work) and social support are the positive effects that may relate to better mental health and well-being.

Indeed, there is evidence to support this model. Adverse psychosocial work characteristics have been associated with increased risk of depressive symptoms (Marchand, Demers, & Durand, 2005; Rugulies, Bültmann, Aust, & Burr, 2006) and common mental disorder (Bültmann, Kant, van den Brandt, & Kasl, 2002; Stansfeld, Bosma, Hemingway, & Marmot, 1998). A recent meta-analysis identified job strain, high demands, and low social support as consistent risk factors for common mental disorder (Stansfeld & Candy, 2006). Job strain has also been associated with increased risk of cardiovascular disease (Bosma et al., 1997; Johnson & Hall, 1988; Schnall et al., 1990).

The Effort Reward Imbalance is an alternative model of the psychosocial work environment (Siegrist, Peter, & Junge, 1990). According to this model, work offers opportunities for people to gain esteem, efficacy, and integration. It is based on the social exchange theory where workers invest effort in work and expect rewards in return. If there is an imbalance between the efforts invested and the reward received then psychological distress occurs with physiological arousal. The effort side of the equation includes both intrinsic effort, which measures people's innate competitiveness, hostility, and involvement with work, and high extrinsic work demands, similar to those measured by Karasek's Model. These efforts are balanced by reward in terms of salary, promotion, or esteem. An imbalance between efforts and rewards has been related to both psychological ill health and physical ill health (Bosma, Peter, Siegrist, & Marmot, 1998; van Vegchel, de Jonge, Bosma, & Schaufeli, 2005).

Other aspects of the work environment have been identified as possibly problematic for health. One involves the concept of organizational justice that has been explored by a Finnish group in a large sample of male and female hospital workers both cross-sectionally and longitudinally in relation to depression (Ylipaavalniemi

et al., 2005). Organizational justice implies that issues at work are dealt with fairly without discrimination and for all employees. Lack of organizational justice has been related to an increased risk of common mental disorders for both men and women (Ybema & van den Bos, 2010). Another important workplace stressor is bullying, usually involving senior employees bullying more junior employees. In a prospective study with female and male hospital employees in Finland, a fourfold increase in risk of common mental disorder over a 2-year period was related to prolonged workplace bullying (Kivimäki et al., 2003). Furthermore, discrimination at work may be related to common mental disorders, and racial or ethnic discrimination has been found to be associated with an increased risk of common mental disorders in a UK studies (Bhui et al., 2005; Wadsworth et al., 2007). Other issues include long work hours and shift work, which have been found to be associated with common mental disorders although the results are not consistent across all studies (Tucker & Knowles, 2008). The association of evening shifts with mental ill health may also suggest a selection factor in that people who are distressed may choose certain shifts.

Criticisms of the Psychosocial Environment Model

Many of the studies that relate psychosocial work characteristics to mental health or coronary heart disease have been criticized because the assessment of work characteristics is based on self-report. Self-report is open to bias from negative affectivity where people have a tendency to report negatively about both their environment and their health. Some studies have overcome this problem by assessing mean values for psychosocial work characteristics related to work groups rather than individuals. Other studies, particularly those in relation to cardiovascular mortality, have classified types of jobs in terms of demands, control, and support and have then related mortality to occupations classified in this way (e.g., Johnson et al., 1989). Other methods to deal with this problem have been to use reports from managers on control and demands in specific posts rather than from individuals, but these have not always been successful as managers may not always be sufficiently informed about the nature of jobs (Stansfeld et al., 1995). A more successful method has been adopted by Virtanen and colleagues (2008) who used an objective measure of work demands in a cohort of hospital workers and related that to sickness absence for depression and prescription of antidepressants in the staff. The objective measure was the percentage of bed occupancy on the ward, which relates directly to the amount of work staff have to carry out and to perceptions of overcrowding and work strain. Thus, although reporting bias may be a feature of associations between the work environment and mental health, using measures that are not susceptible to bias also seems to show the same effects, albeit not as strongly. The observation of a less strong effect might be because bias does indeed contribute to the aforementioned studies of self-reported exposure or because Virtanen's approach does not comprehensively measure environmental stressors.

Occupation and Mental Health

Different occupations are associated with different risks of mental ill health and sickness absence. Many studies have shown fairly consistently high rates for workers in health-related professions, particularly those involved in direct caring, and for teachers in primary and secondary schools, sales personnel, and managers. High rates in these occupations have been shown across Danish, UK, and US studies (Eaton, Anthony, Mandel, & Garrison, 1990; Stansfeld, Rasul, Head, & Singleton, 2011; Wieclaw et al., 2008). However, there are also many variations between studies as to which occupations are most at risk. In many at-risk occupations emotional demands are evident, where people are looking after others but often have to suppress their own emotions. In some occupations low status, low pay, and poor conditions are clearly responsible for increased mental health risk (e.g., kitchen porters). Exposure to violence and public criticism may also be risk factors for mental ill health. Another issue is occupational exposure to alcohol; bar staff had high rates of common mental disorder (29 %) in the 2000 Adult Psychiatric Morbidity Survey presumably partly related to easy access to alcohol, a known risk factor for common mental disorder (Stansfeld et al., 2011). One interesting finding is that occupations within different industry sectors may be more or less at risk. For instance, clerical and secretarial staffs in health-related occupations seem to have lower risk than similar occupations in the construction industry (Stansfeld et al., 2011). There may also be some occupations with dangerous conditions that are preferentially selected by people who have special qualities of resilience and have good mental health.

Selection in Occupations

One possible alternative explanation for higher rates of mental disorders in certain occupations relates to the possibility of selection into jobs according to personality. It is possible that people with health problems, in either themselves or their families, are more likely to work in health-related occupations carrying their preexisting risk for ill health into the occupation. At the other end of the scale, people with qualities of toughness and resilience may be selected into potentially dangerous or risky jobs that they enjoy and such people may be less susceptible to mental health problems. There have been few studies to definitively answer these questions, and it is difficult to separate the effects of selection from those related to the predominant work culture. For instance, men working in predominantly female occupations and vice versa have also been found to have a higher risk for mental disorders (Wieclaw, Agerbo, Mortensen, & Bonde, 2006). In this case, it is difficult to be certain whether high risk relates to selection or to being made to feel an outsider in a particular work culture. Work culture is also particularly important in its own right and may influence the relationship between the employee and the

employer and the ability of the employee to take sickness absence when ill (Johns & Nicholson, 1982). One issue that should not be discounted is the finding that childhood mental ill health predicts that the adults may find themselves in poor working conditions (Clark, Caldwell, Power, & Stansfeld, 2010; Stansfeld, Clark, Caldwell, Rodgers, & Power, 2008, Stansfeld et al., 2008). It seems likely that this represents health selection in which the disturbance of functioning associated with childhood psychological distress may impair education and mean selection into less favorable jobs.

Presenteeism, Productivity, and Sickness Absence

Mental ill health in employees can be expressed in different ways. Employees may, for instance, experience depressive and anxious symptoms but still remain at work. In this case, so-called presenteeism and their functioning and productivity may be affected but the symptoms may not be reported or treated, perhaps for fear of job loss. Alternatively, employees may take time off work with sickness absence, either short spells (usually less than 7 days) or long spells, which need medical certification. Long spells of sickness absence often indicate illness (Marmot, Feeney, Shipley, North, & Syme, 1995), but the determinants of sickness absence are complex. Mental health symptoms predict sickness absence in many studies (Bültmann et al., 2005, 2006). Occupation is also related to sickness absence (Hensing, Andersson, & Brage, 2006). However, recent studies also suggest that the roots of sickness absence may lie in childhood during which absence from school and neurotic symptoms in childhood predict longer term sickness absence and disability (Henderson, Hotopf, & Leon, 2009). Moreover, habits of work learned from parents, coping style and personality, current perceptions of health anxiety and personal vulnerability, and attitudes about work and employment may all contribute to a more complex model of the etiology of sickness absence. Thus, it is no longer satisfactory simply to blame rates of sickness absence on conditions in the workplace.

Increases in Work-Related Stress Disorders

It is interesting that in both the UK (Stansfeld, Clark, et al., 2008, Stansfeld, Woodley Jones, et al., 2008) and Norway rates of sickness absence have increased across the 1990s, possibly decreasing or plateauing slightly since then (Hensing et al., 2006). The increased rates seem to relate particularly to anxiety disorders, neurotic conditions, and depression and not psychoses (Hensing et al., 2006). There is no clear reason why this is the case, but it may be that employees have a greater readiness to report work stress and that mental health conditions are both better recognized and more acceptable as a cause of sickness absence. Although an

increase in stressful working conditions has been proposed, on the whole, this does not seem a plausible explanation as rates of common mental disorders in the general community do not seem to have increased over the same period. Undoubtedly, the nature of jobs has changed in recent years, but whether occupations are generally more stressful is open to question. Certainly, the constantly evolving nature of work and workplaces means that any preventive interventions have to be fit for contemporary working conditions.

Physical Environment of the Workplace

One issue that should be considered is the physical environment of the workplace. The type of ventilation of buildings, exposure to chemicals, access to natural light, and the amount of control employees have over such things as opening windows and where their workstations are placed have all been related to symptoms (Raw, Roys, & Leaman, 1993). There is a considerable literature on sick building syndrome suggesting that certain buildings are bad for health. However, much of this literature does not take into account the psychosocial work environment. A London-based study of civil servants that included both the psychosocial work environment and physical measures of the workplace found that psychosocial variables had a greater influence on symptom measures than did the physical characteristics of the environment, although there were small effects for the ability to control the environment such as the type of ventilation (Marmot et al., 2006).

Vulnerable Groups

A large body of research has documented that some social groups are more vulnerable than others in developing mental health and behavioral problems at work. This research has mainly focused on exposure to experiences of discrimination and stigmatization (Mays & Cochran, 2001). Most of the research on vulnerable groups at work has focused on gender, ethnicity or ethnic minority status, sexuality, mental illness, and learning disability. In particular, research has shown that experiences of discrimination, harassment, and stigmatization can have deleterious effects and lead to great vulnerability to depressive and anxiety disorders, distress, low self-esteem, and alcohol and substance abuse (Mays & Cochran, 2001; Rospenda, Richman, & Shannon, 2009). Negative workplace interactions that are experienced routinely have even been linked to a greater risk of stress-induced mental health problems than major life events (Nawyn, Richman, Rospenda, & Tonda, 2000).

Researchers have focused on specific vulnerable groups to explore how harassment and/or discrimination at work affect mental health and behavioral disorders. There are a lot of similarities between different vulnerable groups on how

harassment and/or discrimination impacts mental health. However, each group faces unique challenges because harassment and/or discrimination is manifested through different facets of one's identity—for example, one's gender in the case of sexual harassment and race or ethnicity in the case of racial harassment (Rospenda & Richman, 2005). In the following section we present a brief review of this research on specific vulnerable groups at the workplace.

Gender

As women's participation in the labor market has been steadily increasing over the past few decades, research on the relationship between women's occupational status in relation to their psychological and mental health has been prolific (McDermott, 2006; Nordenmark, 2002; O'Campo, Eaton, & Muntaner, 2004; Raver, 2003). Sexual harassment exists for men and women and can create a stressful environment at work for both. However, past research has consistently shown that women constitute the great majority of sexual harassment victims (e.g., Rospenda et al., 2009). Indeed, according to the US Equal Employment Opportunity Commission, more than 85 % of the sexual harassment charges across the country were filed by women (Lim & Cortina, 2005).

Numerous studies have documented a plethora of negative outcomes among targets of sexual harassment (e.g., Fitzgerald, Magley, Drasgow, & Waldo, 1999; Rospenda, Richman, Wislar, & Flaherty, 2000). These include job-related consequences (e.g., negative job attitudes, turnover intentions, decreased job satisfaction, organizational commitment), psychological problems (e.g., fear, anxiety, depression), and impaired health (e.g., psychosomatic symptoms) (Lim & Cortina, 2005; Morrow, McElroy, & Phillips, 1994; Ragins & Scandura, 1995). Research has shown that sexual harassment victims frequently suffer from the same types of problems associated with other major workplace stressors. These include increased incidence of psychosomatic illness, reduced concentration, and negative mood (Barling et al., 1996). Prolonged chronic exposure to sexual harassment may cause severe emotional strain and the subsequent disabling symptoms associated with this condition (Dekker & Barling, 1998). Female targets also experience tension, anger, and anxiety, whereas a more limited number experience depression or guilt (Baugh, 1997).

Sexuality

In terms of sexual identity, gay men, lesbians, and bisexual men and women are vulnerable to everyday negative interactions (DiPlacido, 1998), but workplace harassment has been found to be a particularly problematic work stressor for these groups. Workplace harassment can be particularly stressful for gay, lesbian,

and bisexual employees because it can be compounded by (or result from) preexisting expressions of homophobia or perceived vulnerability in one's job due to one's sexual orientation (Nawyn et al., 2000). Studies have reported a high prevalence of verbal insults, threats of violence, and actual violence for gay and lesbian employees compared with their heterosexual counterparts (Norris, 1991) as well as much more sexually harassing behaviors reported by lesbians compared with heterosexual women (Nawyn et al., 2000). An additional workplace stressor for lesbian, gay, and bisexual men and women is related to disclosure of harassing behaviors as studies have found that they rarely report the incident of harassment out of fear of public exposure and, thus, additional harassment (Herek & Glunt, 1993).

The choice of deciding whether to disclose a harassing behavior involves psychological demands that exacerbate stress (Greene, 1994; Woods, 1993). Openness about sexual orientation at the workplace can have an impact at many different levels. It can threaten job security, which is particularly important for lesbians and bisexual women who generally have less financial support from male partners (Dunne, 1997). At its most extreme, lesbians and gay men may be refused a job or may be dismissed (Palmer, 1993) because of their known or perceived sexuality. Less severe, but often equally distressing, is workplace ostracization (Dunne, 1997; Humphrey, 1999), which can take many different forms such as refusing to speak or avoiding contact with a lesbian or a gay colleague, lack of management support for lesbian and gay workers, or refusing to share facilities, such as changing rooms, communal showers, or sleeping accommodation (Skidmore, 2004).

In both US studies and the more limited body of UK research, hostility to homosexuality in the workplace has been found to induce feelings of anxiety, fear, anger, discomfort, worry, and stress (Clarke, 1996; Sorensen & Roberts, 1997), and experiences of occupational heterosexism have been shown to be positively related to depression and psychological distress (Ingram & Smith, 2004; McDermott, 2006). Stress at the workplace can have a significant effect on the general quality of life and satisfaction with life and can lead in many cases to a number of negative behaviors; for example, studies have found that lesbians and gay and bisexual men are more likely than their heterosexual counterparts to engage in excessive alcohol consumption to cope with the stress (Nawyn et al., 2000).

Ethnicity/Race

A large volume of literature on ethnicity and mental health links mental health problems to racial discrimination (Bhui et al., 2005; McKenzie, 2003; Miller & Travers, 2005; Nazroo, 1997, 2003). Racial discrimination is known to be an important risk factor for mental illness in the US (Darity, 2003; Murray, 2003) and UK (Bhui et al., 2005; Wadsworth et al., 2007) workforces. For example, findings from the USA indicate that racial bias, in terms of both institutional and interpersonal prejudice at work, was a more important predictor of job satisfaction

among African American women than other occupational stressors, such as low decision authority and heavy workload (Hughes & Dodge, 1997), and that racial discrimination affects job stress among Black women (Mays, Coleman, & Jackson, 1996). Similarly, research in the UK indicated that almost twice as many non-White respondents reported high or extremely high stress at work compared with White respondents (Smith, Johal, Wadsworth, Davey, & Peters, 2000).

Further research suggests that ethnic minorities are highly exposed to work-related stress (e.g., Miller & Travers, 2005) and that long-term illness in the working population shows significant associations with ethnicity (Census, 2001). Most studies on ethnicity and work link work-related stress to the discrimination experienced at the workplace, which can affect one's sense of control in the work environment as well as workload and opportunities for career development (Bhui et al., 2005). Stress and common mental disorders, which are a result of racial discrimination, can further affect occupational functioning (Bhui et al., 2005). Research also suggests that the interaction between gender and ethnicity can be important when considering the groups that may be more at risk for work-related stress. For example, a well-documented finding is that Black African and Black Caribbean women may be particularly at risk of stress at work (Wadsworth et al., 2007).

Mental Illness

For people with mental health problems, being in paid employment is very important for their recovery. Employment tends to be a significant factor in rehabilitation programs as it has been shown to have a direct relationship with improvements in a number of outcomes such as self-esteem, social support and networks, and psychiatric symptoms (Cook & Razzano, 2000). It is a well-documented finding that people with mental health problems are stigmatized by society and are subjected to discrimination from employers (Stuart, 2006). For example, population surveys have shown that people with mental health problems are significantly more likely to be unemployed than people without mental health problems; in particular, a large-scale population survey found that 61 % of people with mental health problems were unemployed compared with 20 % of the general population (Cook et al., 2005). Unemployment rates also vary by diagnosis, with people suffering from severe mental health problems (e.g., schizophrenia) having the highest unemployment rates (Crowther, Marshall, Bond, & Huxley, 2001).

In addition, people with mental health problems who are in paid employment are less likely to be in prestigious and/or high-earning positions (Chandler, Meisel, Jordan, Rienzi, & Goodwin, 2005) and more likely to be doing unskilled menial jobs and/or be in part-time employment (Rosenheck et al., 2006). Employees with mental health problems may also experience stigma and discrimination from coworkers once their mental illness becomes known. Workers who return to their jobs after an illness report returning to positions of reduced responsibility with

enhanced supervision where they are socially marginalized and become targets for mean-spirited or negative comments from workmates who had previously been supportive and friendly (Cook et al., 2005; Schulze & Angermeyer, 2003). Half of the competitive jobs acquired by people with a serious mental illness will end unsatisfactorily as a result of problems that occur once the job is in progress, largely as a result of interpersonal difficulties (Becker et al., 1998; Stuart, 2006).

Similar to other high-risk groups discussed in this section (e.g., lesbians and gay men, people with learning disabilities), people with mental health problems often find themselves in the stressful situation of whether to disclose their illness (Stuart, 2006). For example, research has shown that people with mental health problems who are employed may not take part in employee assistance programs and may avoid receiving treatment/medication to make sure that their illness is concealed from their colleagues, managers, and employers (Stuart, 2004). Even when people decide to disclose their mental health problems, many organizations (most likely the small ones) will not have sufficient programs in place to address such health problems (Gray, 2000; Harnois & Bagriel, 2000) and managers will not have the knowledge or the skills to effectively manage mental health problems at the workplace (Stuart, 2006).

The relationship between work and psychosis such as schizophrenia differs from that between work and common mental disorders. Many people with psychosis are severely disabled with both positive symptoms—persisting delusions and hallucinations—and often more disabling, negative symptoms such as apathy, neglect, and social withdrawal that interfere with the ability to work. For people with chronic schizophrenia there are many benefits of working in relation to social status, structure of the day, availability of social networks, and self-esteem (Velligan & Gonzalez, 2007). Regrettably there are sometimes perverse incentives that mean that potential employees have a financial incentive to remain on benefits rather than to return to work. There has been considerable success in getting unemployed people with serious mental illness back into work through the individual placement and support program that supports both the patient and the employer in placing the person with serious mental illness in open employment (Burns et al., 2007). The disadvantage of these jobs is that many are entry-level or part-time jobs where the work characteristics are fairly poor. Nevertheless, working in such jobs may be better than no job at all because lack of activity and poverty of environment can be particularly damaging to people with chronic psychotic illnesses (Wing, 1970). Because there has been fairly little research on the effect of working conditions on people with serious mental illness, this will not be a major focus of the chapter.

Learning Disabilities

A great deal of what has been discussed in relation to people with mental health problems applies to people with learning disabilities (LDs) in terms of disclosure, type of job, and earnings. A further contributing factor is that a significant

proportion of people with LDs suffer from mental health problems (Smiley, 2005), which is believed to partly stem from the socioeconomic disadvantage and related social exclusion that people with LDs tend to experience. In particular, according to figures from the Office for National Statistics in the UK, only 17 % of working-age people with learning difficulties had a paying job compared with 67 % of men and 53 % of women without LDs (Emerson & Hatton, 2007). For those people with LDs in paid employment, research has shown that they are more likely to be in part-time employment and earn a lot less than people without LDs (Madaus, Foley, McGuire, & Ruban, 2002).

Although people with LDs are suggested to be one of the most vulnerable groups at the workplace, research on their experiences at work and how these experiences affect mental health is scarce. The little research that exists focuses mainly on whether people with LDs find satisfying employment as well as issues of disclosure. Self-disclosure is an intensely personal and complex decision, as the costs and benefits of disclosure must be weighed against a variety of concerns (Baldridge & Veiga, 2001; Gerber & Price, 2003; Hughes & Williams, 1994). A study found that although 80 % indicated that their LDs affected their work in some way, only 20 % self-disclosed their LD status during the interview process and 43 % self-disclosed after they were hired (Madaus, 2008).

Age

Although the term "ageism" was used to refer to the process of stereotyping and discrimination against people perceived as too old, it has now come to refer more broadly to prejudice and discrimination on the grounds of either being too old or too young (e.g., Palmore, 1990). For example, a survey of undergraduate students showed that over one-third of the participants reported experiences of age-related discrimination (Duncan & Loretto, 2004; Loretto, Duncan, & White, 2000).

Most studies on the impact of age discrimination on the employee have mainly focused on career rather than on mental health. In particular, studies of performance appraisal report older employees receive lower performance ratings than their younger counterparts (Saks & Waldman, 1998) and they are more likely than their younger counterparts to experience discrimination in access to training and development opportunities (Department of Employment, 1994) as well as promotion opportunities (Palmore, 1990). Older workers are also more likely to be selected for redundancy than are middle-age workers (Snape & Redman, 2003). Young people who by definition have little experience and skills face great difficulties in finding the right employment and are often underemployed (Barling & Kelloway, 1999; Lindley, 1996), seen as untrustworthy, and given less responsibility (Loretto et al., 2000).

The small number of studies that explored the impact of ageism on an employee's economic and psychological well-being has largely focused on older employees. For example, on leaving the workforce, older workers find it much more

difficult to reenter it (Sargeant, 2001) and experience longer periods of unemployment between jobs than their younger counterparts (McGregor & Gray, 2001). Greater prospects of redundancy are also likely to be particularly stressful for older workers if there is little likelihood of obtaining meaningful work ever again (Wood, Wilkinson, & Harcourt, 2008). Furthermore, those who feel that they have been discriminated against because of their age may feel under-rewarded in terms of pay and status, while discriminatory treatment in general may create a sense of being undervalued by the organization (Hassell & Perrewe, 1993), which can lead to mental health problems such as negative affective states (Gutek, Cohen &, Tsui, 1996) and stress (Snape & Redman, 2003). Moreover, those experiencing discrimination will feel less loyal and have a lower level of identification with the organization and will feel unsupported (Snape & Redman, 2003).

Conclusion

Harassment and discrimination in the workplace take many forms and can occur for a variety of reasons such as one's gender, sexuality, race/ethnicity, or age. Furthermore, these vulnerability factors can interact. For example, age discrimination can incorporate a gender dimension, with older women being particularly negatively stereotyped (Duncan & Loretto, 2004; Schuman & Kleiner, 2001; Walker, 2005; Wood et al., 2008). Evandrou and Glaser (2004) note that, having fulfilled child-raising and other family social care obligations, older women are very much more likely to face poverty in their old age. Research has also found that African Caribbean women may be particularly at risk of work stress and its consequences because of the potential for gender and/or racial bias that may have been apparent for sometime (Wadsworth et al., 2007). While gender, race/ethnicity, and age may be at interplay, other sources of powerlessness (e.g., sexual orientation) merit future attention. For example, although sexuality is not the same as gender, it cannot be understood without reference to it in the context of organizations (Skidmore, 2004). Therefore, multiple social statuses, such as gender, sexuality, age, race/ethnicity, education, and income, may influence cumulatively specific psychiatric vulnerabilities among individuals at the workplace in ways that are not yet understood.

Interventions in the Workplace

We have already outlined theoretical models for workplace effects on mental health. It should follow that workplace interventions could be designed to prevent, screen for, or treat common mental disorders, as well as promote well-being.

Indeed, from a public health point of view, the workplace is logistically attractive for the delivery of preventive interventions in mental health, given its established infrastructure and the numbers of adults who are located there.

Moreover, it is a potentially advantageous corporate strategy; research has linked depression and other mental health problems to impaired work performance, absenteeism, and premature retirement (Haslam, Atkinson, Brown, & Haslam, 2005; Sanderson & Andrews, 2006). That said, to actually make changes to improve the workplace environment can be challenging for a variety of reasons. A separate challenge is to prove the effectiveness of any specific intervention as systematically and rigorously as possible.

In this section, types of workplace interventions are categorized and related to the models of the psychosocial work environment as outlined earlier in this chapter. We then turn to studies on workplace interventions, particularly on interventions that target the workplace environment itself, for an assessment of their efficacy for mental health. Gaps in the literature, as well as the limitations of workplace intervention research, are identified, and recommendations are made regarding the future direction that research should take.

Categorizing Intervention Strategies

Preventive interventions can be categorized into primary, secondary, and tertiary strategies. These are differentiated by the timing of their implementation in the disease process (Couser, 2008). In the context of mental health in the workplace, primary prevention is the universal application of an intervention to all individuals within a defined workplace population to protect them from mental ill health; secondary prevention involves intervention after a process of selecting/screening for those with risk factors for mental health problems; and tertiary prevention are measures to treat and reduce the disability related to mental health problems of those who have already developed them.

Intervention strategies may also be categorized according to the organizational level at which the interventions are made. In relation to the workplace, they may be located at the legislative or policy level (i.e., outside the workplace itself), at the employer level, at the level of the job/task occupied by individuals, and at the level of individuals. Specific strategies for primary, secondary, and tertiary workplace interventions at these four different levels have been provided by the National Institute for Occupational Safety and Health (NIOSH), with additions made by Landsbergis (2009) (Table 10.1). For example, Europe-wide initiatives at the legislative level have included a European Union directive in 1989 to "alleviate monotonous work at predetermined pace," a European Commission Guidance document on work-related stress in 2000, and an agreement by major employers and unions in Europe in 2004 that obliges employers to protect employees from work-related stress. Interventions at the employer level tend to target human resources policies and the working environment by making it more supportive. Interventions at the task/job level, as well as the employer level, have been responsible for introducing increased job control and participation in decision-making, job enrichment, and job redesign, whereas interventions at the individual

Table 10.1 Levels of work organization interventions and stages of prevention

Levels of intervention	Primary	Secondary	Tertiary	Typical intervention methods
Legislative/ policy	Work hour limits, *Ban mandatory overtime, staffing requirements*	Workers compensation	Social Security Disability	*Legislation, regulation*
Employer/ organization	Work–family programs; *Workplace health and safety programs; New systems of work organization; Work-site surveillance*	Health promotion programs; *Work-site surveillance*	Company provided long-term disability; Return to work programs	*Collective bargaining; Worker ownership; Employer-initiated programs/policies, including contingent work, downsizing, new systems of work organization*
Job/task characteristics	Job/task redesign; Job rotation; *Integrated workplace health-health promotion programs*	Provision of light duty jobs		*Labor-management committees, initiatives; Action research; Employer-initiated job redesign*
Individual/ job interface	Health promotion programs	Stress management programs Employee assistance programs Disease management programs		*Employer-initiated programs; Labor-management initiatives*

Additions to NIOSH's basic model by Landsbergis in italics

level often involve targeted programs such as health promotion programs (Landsbergis, 2009). Of course, any one prevention strategy may involve interventions at several levels, each mutually reinforcing the other. Moreover, it may be more effective to intervene at both organizational and individual levels simultaneously.

Designing Intervention Strategies

Given this available range of intervention strategies, which one should an organization choose to improve its working environment?

Where the workplace itself is likely to generate widespread mental illness risks, it may be more efficacious to adopt primary intervention strategies. These would target the occupational- and employment-related factors responsible for increasing risk (e.g., taking steps to reduce the risk of exposure to violence in staff dealing with a potentially hostile client group). Primary prevention strategies are applied across the entirety of a defined population and, hence, could have far-reaching benefits on workforce morale. However, given their universal target, they might be thinly applied. Therefore, where risk is restricted to a few and appears in clusters, primary prevention strategies could waste resources on those who don't need help, while not doing enough to help those who do need it.

In such a situation, it may be more efficacious to adopt a secondary prevention strategy, particularly where impact of the workplace adversity is strong and at-risk groups readily identifiable. These strategies might tackle discrimination in high-risk social groups, as discussed earlier in this chapter. However, a secondary prevention strategy can come with difficulties: it may in itself be perceived as discriminatory, invite feelings of victimization, and be ultimately counterproductive. Secondary intervention strategies must, therefore, be carefully designed to best maintain privacy and dignity.

The efficacy of interventions will also depend on whether they correctly identify the most pertinent problems and succeed in addressing them. Theories such as the Demand–Control Model, the Iso-Strain Model, the Effort Reward Imbalance Model, and "organizational justice" play a crucial role in this regard, in that they provide models against which the psychosocial environment of any workplace may be assessed and improved. Employer- and job-level approaches are theoretically best suited to make changes based on these models and may be most effective in practice, especially if sustainable improvements in working conditions are to be achieved.

Lastly, mental health problems are associated with some occupations more than others. As already discussed, this may be because individuals self-select into occupations according to their personality and predilections or may be a true effect of the job. So, for example, although the highest rates of mortality from alcohol abuse are found among bar workers and publicans (Romeri, Baker, & Griffiths, 2007), this may be due to the predisposition of individuals attracted to the industry

rather than the opportunities for alcohol consumption that these occupations offer. Explanations for the high rates of substance abuse among anaesthesiologists compared with other medical specialists may be subject to a similar debate (Bryson & Silverstein, 2008). Whether a primary or secondary intervention strategy is adopted will largely depend on the strength of the evidence supporting the "opportunities" model, which invites a primary intervention strategy at the organizational/occupational level, versus the "self-selection" model, which invites a secondary intervention strategy at the individual level.

Indeed, the success or failure of interventions that have been devised under the influence of specific theoretical models may themselves be employed to lend or reduce support for the models by which they have been shaped. The questions that now concern us are what evidence exists for the success and validity of recent interventions that target the workplace? How useful are they for promoting our understanding of the links among psychosocial aspects of the work environment, mental well-being, and mental health problems? And how can they inform future policy and practice in the workplace?

The Use of Meta-analysis

Firstly, is there evidence to suggest that workplace interventions can improve mental health? To answer this question, it may be useful to consider the findings of meta-analyses that have reviewed this literature. Because they pool data from a large number of studies, meta-analyses may provide us with shortcut answers to questions of effectiveness. And because they select only those studies that meet specified scientific criteria, they may also narrow the field to those studies that are methodologically robust.

Most meta-analyses of the recent literature on interventions in the workplace do indeed find a small but statistically significant overall benefit of workplace interventions on outcome measures such as absenteeism, psychological distress, and anxiety/depression scales (Table 10.2). These meta-analyses focused on a variety of intervention strategies. What is disappointing, however, is that most meta-analyses include very few studies that directly target employer-/job-level risk factors (e.g., poor psychosocial work environment), which is the main topic of this chapter and probably the most important causal link between the workplace and poor mental health. For example, in a meta-analysis of 48 studies of interventions to reduce occupational stress between 1977 and 1996, van der Klink and colleagues (2001) found only five studies that were targeted at the employer/job level. Instead, by depending on changes at an individual level (such as stress management or cognitive behavioral therapy [CBT]), these studies are merely using the workplace to roll out healthcare interventions and not intervening in the workplace environment itself. Why is this happening? The lack of organizational interventions in meta-analyses presumably reflects the difficulty in organizing randomized and controlled versions of employer-level trials (Gardell & Gustavsen, 1980), which are therefore

Table 10.2 Meta-analyses on workplace interventions for mental health

Author (search dates)	Type of intervention	Inclusion criteria	Organizational level	Outcome measures	Result
Martin, Sanderson, & Cocker, 2009 (1997–2007)	Health promotion interventions (e.g., stress management programs, exercise programs)	Primary and/or secondary interventions. RCT and quasi-experimental.	Individual (14) and/or employer (3)	Standardized screening measure for depression and/or anxiety	Small but significant benefit for depression (SMD 0.28; 95 % CI 0.12–0.44) and anxiety (SMD 0.29; 95 % CI 0.06–0.51 for anxiety)
Parks & Steelman, 2008 (1980–2005)	Organizational wellness programs (e.g., fitness program, stress-reduction, nutrition)	Primary interventions. Any empirical study.	Individual (17)	Absenteeism; job satisfaction	Small benefit for absenteeism (SMD 0.3; 95 % CI 0.48–0.22) and moderate for job satisfaction (SMD 0.42; 95 % CI 0.05–0.8)
Marine, Ruotsalainen, Serra, & Verbeek, 2006 (up to 2005)	Individual-level (e.g., CBT, relaxation); employer-level (e.g., training in changing work attitudes, training in support mobilization, schemes to change task structure)	Primary interventions. Healthcare workers only. RCTs and quasi-experimental studies.	Individual (13) and/or employer (6)	Questionnaire on stress and burnout; anxiety/depression symptoms	Individual-level studies small but significant reduction in stress, and anxiety symptoms; employer-level studies small but significant reduction in stress (SMD 0.34; 95 % CI 0.62–0.06)
van der Klink, Blonk, Schene, & van Dijk, 2001 (1977–1996)	CBT; relaxation programs; multimodal programs; employer-based interventions	Primary interventions. RCT or quasi-experimental.	Individual (43) and/or employer (5)	Anxiety and depression symptoms; absenteeism	Small but significant benefit (SMD 0.44; 95 % CI 0.36–0.52); employer-level intervention studies showed no benefit (SMD 0.08; 95 % CI–0.03–0.19)

Saunders, Driskell, Johnston, & Salas, 1996 (1977–1991)	Stress inoculation training (CBT approach to stress management) in high-anxiety jobs	Primary interventions. RCT and quasi-experimental.	Individual (37)	Performance anxiety; state anxiety; performance under stress	Strong reduction of performance anxiety (SMD 0.509; $p < 0.01$); moderate reduction of state of anxiety (SMD 0.373; $p < 0.01$); and enhanced performance (SMD 0.296; $p < 0.01$)

being omitted from meta-analyses in the pursuit of scientific rigor. One reviewer, for example, found them "difficult to interpret, causally ambiguous, inconsistent, based on small samples and/or statistically insignificant" (Parkes & Sparkes, 1998).

If meta-analyses cannot provide shortcut answers as to the efficacy of interventions at the employer/job level, we must instead look to data from individual studies.

Evaluating Organizational Interventions: Action Research

One promising organizational approach has involved and empowered those in the workplace to identify problems in, and emanating from, the workplace, as well as involved them in the development of interventions to change these problems and measure their effects. Referred to commonly as action research, it has been common in Scandinavia (International Labour Organization, 1992). For example, the involvement of bus drivers in action research to reduce stress and improve their health was responsible for identifying and changing several aspects of job design, such as lanes and routes, signal priority, and bus stops. This led to a significant decline in reported "job hassles per hour" as well as systolic blood pressure (Poulsen, Jensen, Bach, & Schostak, 2007). Indeed, a number of reviews have suggested that the most promising interventions at the organizational level are those that have included the meaningful participation of employees (LaMotagne, Keegel, Louie, Ostry, & Landsbergis, 2007; Parker, Jackson, Sprigg, & Whybrow, 1998; Semmer, 2006).

This approach might be seen to act through Karasek's Job Control Model, since control, demand, and/or support could theoretically be improved. Egan and colleagues (2007) performed a systematic review of action research studies to test Karasek's Model. Of 18 identified studies, nearly all evaluated interventions involving "problem-solving" or "participatory" committees of employee representatives that aimed to tackle workplace stressors. Some studies combined this with individual-level interventions, such as relaxation techniques. One should note that no studies were randomized, and many contained non-matched groups for comparison.

Nevertheless, it is interesting that of the eight studies to report overall post-intervention improvement in job control/participation, seven also reported significant overall health improvements (such as improvement in mental health questionnaire scores). Four studies reported decreased job demands after the intervention and were accompanied by improved health outcome in each instance. Improved support was also associated with improved health in the majority of studies in which it was measured. In those studies where control, demand, or support were recorded as unchanged or worsened, health outcomes often remained unchanged (Egan et al., 2007).

Collectively, these observations go some way to broadly supporting the individual components—control, demand, and support—of Karasek's Job Control Model.

However, no one study within the systematic review measured and improved on all three components in a meaningful way, and no studies correlated the association between these components and health outcomes within individuals. To follow Karasek's Model more explicitly might benefit the design and success of future studies.

It is also worth noting that only two studies stratified outcomes by "at risk groups." Those that did suggested that lower grade workers and employees from ethnic minorities might benefit most from interventions (Park, Schaffer, & Griffin-Blake, 2004; Parker, Chmiel, & Wall, 1997). Future studies should aim to gather such data, as they could support a role for interventions in decreasing healthcare inequalities.

Workplace Reorganization Strategies

To make a large and lasting improvement on mental health in the workplace, it might be that interventions have to target fundamental workplace reorganization.

For example, there has been a long tradition of workplace research that examined the impact of assembly line production, the dominant mode of work organization of the early twentieth century that is often referred to as "Fordism" or "Taylorism," and reported findings of burnout, stress-related illness, alienation, and absenteeism (Landsbergis, Cahill, & Schnall, 1999). As a result, new systems of work organization have been introduced to replace the assembly line. However, evaluation of the effects of these new systems has taken various forms, appears to depend on political circumstances and the power of the labor movement, and frequently avoids examination of mental health effects altogether. For example, the Japanese lean production approach, which was devised to replace the assembly line, has been evaluated more in terms of the interests of management and shareholders, such as productivity, profitability, and product quality (Landsbergis et al., 1999) and less for its effects on employees.

In Sweden and Norway, by contrast, changes to the organization of work have not only been evaluated in terms of their health outcomes, but the interventions have been designed with this in mind. Research in the 1960s and 1970s had demonstrated the impact of the assembly line, with its monotonous and repetitive machine-paced work, on mental and physical health (Gardell & Gustavsen, 1980). Evaluations of work–life programs in Sweden and Norway reportedly provide evidence of the impact of increased job control on work and life satisfaction, as well as improvements in productivity (Levi, 2000). The work humanization interventions in the late 1980s and early 1990s at Saab and Volvo may be singled out for special mention. Instead of the 60-s cycle times of the typical assembly line, parallel teams built complete cars with cycle times of 1.5–7 h. Unfortunately, the closure of the plant prevented adequate evaluation of the parallel teams experiment for its impact on chronic illnesses and absenteeism.

Further workplace reorganization interventions have been evaluated for their impact on mental health. Bambra and colleagues (2007) performed a systematic review of workplace reorganization studies and found 19 studies with control/comparison groups, although no randomized studies. Eight of these studies aimed to introduce more job-task variety, for example, by allowing healthcare workers or production line workers more breadth in their tasks and responsibilities. Seven different studies aimed to promote teamwork and increase the collective responsibility and decision-making power of the employees. Six studies aimed to increase autonomy and collective involvement, for example, by using the "lean production" or "just-in-time" models of mass production rather than traditional assembly lines. Studies targeted diverse groups including factory workers, doctors, carers, and office workers (Bambra, Egan, Thomas, Petticrew, & Whitehead, 2007).

Results were mixed, with many studies reporting no benefit from the intervention. For example, the job-task variety interventions often failed to increase job control and job demand often increased; mental health benefits were small or not significant at best. The team working interventions did improve the work environment, by increasing support, but did little for health. The autonomous work groups actually worsened the psychosocial work environment, contrary to their aims, with decreased control and autonomy; job anxiety and depression increased in several cases (Bambra et al., 2007).

It is difficult to generalize as to the implications of these studies, because they differ markedly in terms of aim, design, intervention, and outcome measures, and many had non-matched comparison groups. Further, much of the primary literature does not specify exactly what changes were made in specific terms. One should note that many did not actually succeed in changing the psychosocial work environment, nor did they aim to (focusing instead on improving productivity). These economic-driven trials tended to cause negative or negligible health outcomes. One interesting point is that "job control" seemed to influence mental health more than "job demand" and "job support," suggesting that this might be the most important component in Karasek's Model in practice although demands have often been found to be more important in observational studies.

Other research has attempted to improve the physical environment, including technological improvements (Kawakami, Araki, & Kawahima, 1997) and advising on posture and lifting (Evanoff, Bohr, & Wolf, 1999). The former study showed no improvement in psychosocial environment, perhaps because the intervention was coupled with an increase in workload. The latter study did report improved musculoskeletal health and reduced absenteeism but did not report mental health outcomes.

Challenges and Future Directions

Evaluations of employer-level interventions that target the effects of workplace stressors on mental health are limited. Where they exist, the evidence has not always been positive and reviewers have found the evidence to be vague, weak, and contradictory (Parkes & Sparkes, 1998). Is this a challenge that can be met?

Firstly, it is clear that information to identify the locations and sources of high risk in the workplace must be gathered more systematically. "Work site surveillance," as called for in the Tokyo Declaration of 1998, is required both to identify clusters of high risk and to measure work-related stressors and the health problems that result from them (Landsbergis, 2009).

Secondly, study designs must be adjusted to ensure that longer timeframes are included. Evaluations of interventions at the individual level are frequently limited in value by the absence of follow-ups (Martin et al., 2009). While a number of such studies have demonstrated substantial impact, questions have been raised as to the sustainability of these interventions, given the unchanged nature of the workplace itself. In contrast, the effects of interventions at the organizational level may appear weak because they take longer to manifest themselves (Martin et al., 2009; van der Klink et al., 2001). One of the challenges facing the researcher, therefore, is to successfully impress on funders and sponsors the long-term nature of this particular kind of evaluation project. In a climate that demands quick results, this may be more difficult than imagined.

Nevertheless, some initial steps may be taken. If researchers are to promote organizational interventions in the workplace and encourage the replication of evaluations with positive outcomes, they must provide the basic information to make replication possible. Many of the studies that have evaluated organizational interventions have lacked information about the practical details that permit replication, such as details on the nature of the interventions (their content), how they were applied (their process), and their cost to the organization. Promising interventions will only be disseminated, and their evaluation replicated, if potential sponsors and decision makers are provided with careful and detailed documentation of effective intervention initiatives.

Difficulties associated with evaluating interventions are compounded by uncertainties over the nature of the findings and what they represent. This is particularly so in the area of mental health and mental illness, where indicators of health and illness are subject to considerable debate. Hence, clarity about the nature of the effects achieved should be central to any evaluation, as should the enhancement of methodological rigor.

This is no mean task. Many of the studies reported, including interventions targeted at the individual level, rely on subjective rating skills. Some of the measures taken to deal with reporting bias have already been referred to in relation to establishing the links between the psychosocial work environment and risks to physical and mental ill health. However, additional reasons make it difficult for evaluations of interventions at the organizational level to achieve methodological

rigor (Kompier, 2003). Firstly, controlled trials are possible, but there are practical difficulties associated with them. Unless there are many groups/plants/offices/work locations that are similar, with some chosen for the intervention by scientific random sampling methods and the others remaining as controls, randomization of the "population" of work sites for intervention is very difficult. Nevertheless, this difficulty is not insurmountable, as exemplified by the WellWorks Project, which conducted a randomized controlled trial on cancer prevention strategies in 24 work sites in Massachusetts (Sorensen, Stoddard, Ockene, Hunt, & Youngstrom, 1996).

Secondly, by virtue of the absence of "blinding," interventions in the workplace are subject to the same difficulties that accompany all interventions in the social world. In any social intervention study, those subject to the intervention are likely to be aware that some changes have been made to their environment. This can have several consequences for workplace interventions that evaluate for changes in mental health. Because of the intervention, subjects may convince themselves that there have been improvements in their well-being, whether or not this is the case. Alternatively, there may be actual improvements in mental health. However, well-being may have been promoted not because of the substance of the intervention itself, but because of the fact that there was any intervention at all. Often referred to as the Hawthorne effect, after research carried out by Mayo and colleagues in the 1920s and 1930, the awareness of participants that they are subject to an intervention may "tamper" with the results of any social experiment (Landsberger, 1958; Mayo, 1949).

However, the lack of studies that evaluate interventions at the employer level is not only due to the aforementioned methodological challenges. Interventions targeted at individuals appear to be influenced by prevailing political philosophies and associated with the weakening power of labor. Only this can explain such radically different approaches to workplace interventions in Japan, the USA, and Scandinavia. Where individualism dominates, structural causes of ill health (whether mental or physical) are viewed with suspicion whereas lifestyle choices are perceived to be the key to well-being and ill health. Under these circumstances, the most crucial challenge facing researchers may be not only developing robust methods for evaluating organizational interventions but also ensuring that the topic finds itself on the research and development agenda in the first place.

References

Baldridge, D. C., & Veiga, J. F. (2001). Toward a greater understanding of the willingness to request an accommodation: Can requesters' beliefs disable the Americans with Disabilities Act? *Academy of Management Review, 26*, 85–99.

Bambra, C., Egan, M., Thomas, S., Petticrew, M., & Whitehead, M. (2007). The psychosocial and health effects of workplace reorganisation: 2, A systematic review of task restructuring interventions. *Journal of Epidemiology and Community Health, 61*(12), 1028–1037.

Barling, J., Dekker, I., Loughlin, C., Kelloway, K., Fullager, C., & Johnson, D. (1996). Prediction and replication of the organisational and personal consequences of workplace sexual harassment. *Journal of Managerial Psychology, 11*(5), 4–25.

Barling, J., & Kelloway, E. K. (1999). *Young workers: Varieties of experiences*. Washington, DC: American Psychological Association.

Bartley, M. (1997). Unemployment and ill health: Understanding the relationship. *Journal of Epidemiology and Community Health, 48*(4), 333–337.

Baugh, S. G. (1997). On the persistence of sexual harassment in the workplace. *Journal of Business Ethics, 16*, 899–908.

Becker, D. R., Drake, R. E., Bond, G. R., Xie, H., Dain, B. J., & Harrison, K. (1998). Job terminations among persons with severe mental illness participating in supported employment. *Community Mental Health Journal, 34*, 71–82.

Bhui, K., Stansfeld, S., McKenzie, K., Karlsen, S., Nazroo, J., & Weich, S. (2005). Racial/ethnic discrimination and common mental disorders among workers: Findings from the EMPIRIC Study of Ethnic Minority Groups in the United Kingdom. *American Journal of Public Health, 95*, 496–501.

Bosma, H., Marmot, M. G., Hemingway, H., Nicholson, A. C., Brunner, E., & Stansfeld, S. A. (1997). Low job control and risk of coronary heart disease in Whitehall II (prospective cohort) study. *British Medical Journal, 314*(7080), 558–565.

Bosma, H., Peter, R., Siegrist, J., & Marmot, M. (1998). Two alternative job stress models and the risk of coronary heart disease. *American Journal of Public Health, 88*(1), 68–74.

Bryson, E. O., & Silverstein, J. H. (2008). Addiction and substance abuse in anesthesiology. *Anesthesiology, 109*(5), 905–917.

Bültmann, U., Huibers, M. J., van Amelsvoort, L. P., Kant, I., Kasl, S. V., & Swaen, G. M. (2005). Psychological distress, fatigue and long-term sickness absence: Prospective results from the Maastricht Cohort Study. *Journal of Occupational and Environmental Medicine, 47*(9), 941–947.

Bültmann, U., Kant, I. J., van den Brandt, P. A., & Kasl, S. V. (2002). Psychosocial work characteristics as risk factors for the onset of fatigue and psychological distress: Prospective results from the Maastricht Cohort Study. *Psychological Medicine, 32*, 333–345.

Bültmann, U., Rugulies, R., Lund, T., Christensen, K. B., Labriola, M., & Burr, H. (2006). Depressive symptoms and the risk of long-term sickness absence: A prospective study among 4747 employees in Denmark. *Social Psychiatry and Psychiatric Epidemiology, 41*(11), 875–880.

Burns, T., Catty, J., Becker, T., Drake, R. E., Fioritti, A., Knapp, M., et al. (2007). The effectiveness of supported employment for people with severe mental illness: A randomised controlled trial. *Lancet, 370*(9593), 114652.

Census. (2001). London: Office for National Statistics.

Chandler, D., Meisel, J., Jordan, P., Rienzi, B. M., & Goodwin, S. N. (2005). Mental health, employment, and welfare tenure. *Journal of Community Psychology, 33*, 587–609.

Clark, C., Caldwell, T., Power, C., & Stansfeld, S. A. (2010). Does the influence of childhood adversity on psychopathology persist across the life course? A 45-year prospective epidemiologic study. *Annals of Epidemiology, 20*(5), 385–394.

Clarke, G. (1996). Conforming and contesting with (a) difference: How lesbian students and teachers manage their identities. *International Studies in Sociology of Education, 6*(2), 191–209.

Cook, J. A., Lehman, A. F., Drake, R., McFarlane, W. R., Gold, P. B., & Leff, H. S. (2005). Integration of psychiatric and vocational services: A multisite randomized, controlled trial of supported employment. *American Journal of Psychiatry, 162*, 1948–1956.

Cook, A., & Razzano, L. (2000). Vocational rehabilitation for persons with schizophrenia: Recent research and implications for practice. *Schizophrenia Bulletin, 26*, 87–103.

Couser, G. P. (2008). Challenges and opportunities for preventing depression in the workplace: A review of the evidence supporting workplace factors and Interventions. *Journal of Occupational and Environmental Medicine, 50*, 411–427.

Crowther, R. E., Marshall, M., Bond, G. R., & Huxley, P. (2001). Helping people with severe mental illness to obtain work: Systematic review. *British Medical Journal, 322*, 204–208.

Darity, W. A., Jr. (2003). Employment discrimination, segregation, and health. *American Journal of Public Health, 93*, 226–231.

Dekker, I., & Barling, J. (1998). Personal and organizational predictors of workplace sexual harassment of women by men. *Journal of Occupational Health Psychology, 3*(1), 7–18.

Department of Employment. (1994). *Training statistics*. London: HMSO.

DiPlacido, J. (1998). Minority stress among lesbians, gay men, and bisexuals: A consequence of heterosexism, homophobia, and stigmatization. In G. Herek (Ed.), *Stigma and sexual orientation: Understanding prejudice against lesbians, gay men, and bisexuals (Psychological Perspectives on Lesbian and Gay Issues)* (Vol. 4, pp. 138–159). Thousand Oaks, CA: Sage Publications.

Duncan, C., & Loretto, W. (2004). Never the right age: Gender and age-based discrimination in employment. *Gender, Work and Organization, 11*(1), 95–115.

Dunne, G. A. (1997). *Lesbian lifestyles: Women's work and the politics of sexuality*. Toronto, ON: University of Toronto Press.

Eaton, W. W., Anthony, J. C., Mandel, W., & Garrison, R. (1990). Occupations and the prevalence of major depressive disorder. *Journal of Occupational Medicine, 32*, 1079–1087.

Egan, M., Bambra, C., Thomas, S., Petticrew, M., Whitehead, M., & Thomson, H. (2007). The psychosocial and health effects of workplace reorganisation. 1. A systematic review of organisational-level interventions that aim to increase employee control. *Journal of Epidemiology and Community Health, 61*(11), 945–954.

Emerson, E., & Hatton, C. (2007). Mental health of children and adolescents with intellectual disabilities in Britain. *British Journal of Psychiatry, 191*, 493–499.

Evandrou, M., & Glaser, K. (2004). Family, work and quality of life. *Ageing and Society, 24*, 771–791.

Evanoff, B. A., Bohr, P. C., & Wolf, L. D. (1999). Effects of a participatory ergonomics team among hospital orderlies. *American Journal of Industrial Medicine, 35*, 358–365.

Fitzgerald, L. F., Magley, V. J., Drasgow, F., & Waldo, C. R. (1999). Measuring sexual harassment in the military: The sexual experiences questionnaire (SEQ-DoD). *Military Psychology, 11*, 243–263.

Fryers, T., Melzer, D., & Jenkins, R. (2003). Social inequalities and the common mental disorders: A systematic review of the evidence. *Social Psychiatry and Psychiatric Epidemiology, 35*(5), 229–237.

Gardell, B., & Gustavsen, B. (1980). Work environment research and social change: Current developments in Scandinavia. *Journal of Occupational Behavior, 1*, 3–17.

Gerber, P. J., & Price, L. A. (2003). Persons with learning disabilities in the workplace: What we know so far in the Americans with Disabilities Act era. *Learning Disabilities Research and Practice, 18*, 132–136.

Gray, P. (2000). *Mental health in the workplace: Tackling the effects of stress*. London: The Mental Health Foundation.

Greene, B. (1994). Lesbian and gay sexual orientations: Implications for clinical training, practice, and research. In B. Greene & G. M. Herek (Eds.), *Lesbian and gay psychology: Theory, research, and clinical applications* (pp. 1–24). Thousand Oaks, CA: Sage Publications.

Gutek, B. A., Cohen, A. G., & Tsui, A. (1996). Reactions to perceived sex discrimination. *Human Relations, 49*(6), 791–813.

Harnois, G., & Bagriel, P. (2000). *Mental health and work: Impact, issues and good practices*. Geneva: World Health Organization and International Labour Organization.

Haslam, C., Atkinson, S., Brown, S., & Haslam, R. (2005). Anxiety and depression in the workplace: Effects on the individual and organisation. *Journal of Affective Disorders, 88*(2), 209–215.

Hassell, B. L., & Perrewe, P. L. (1993). An examination of the relationship between older workers' perceptions of age discrimination and employee psychological states. *Journal of Managerial Issues, 5*(1), 109–120.

Henderson, M., Hotopf, M., & Leon, D. A. (2009). Childhood temperament and long-term sickness absence in adult life. *British Journal of Psychiatry, 194*(3), 220–223.

Hensing, G., Andersson, L., & Brage, S. (2006). Increase in sickness absence with psychiatric diagnosis in Norway: A general population-based epidemiologic study of age, gender and regional distribution. *BMC Medicine, 22*(4), 19.

Herek, G. M., & Glunt, E. K. (1993). Interpersonal contact and heterosexuals attitudes toward gay men: Results from a national survey. *The Social Sex Research, 30*(3), 239–244.

Hughes, D., & Dodge, M. A. (1997). African American women in the workplace: Relationships between job conditions, racial bias at work and perceived job quality. *American Journal of Community Psychology, 25*, 581–599.

Hughes, C., & Williams, G. J. (1994). To disclose or not to disclose: That is the question. *Transition LINC, 2*, 1–6.

Humphrey, J. (1999). Organizing sexualities, organized inequalities: Lesbians and gay men in public service occupations. *Gender, Work and Organization, 6*(3), 134–351.

Ingram, K. M., & Smith, N. G. (2004). Workplace heterosexism and adjustment among lesbian, gay, and bisexual individuals: The role of unsupportive social interactions. *Journal of Counselling Psychology, 51*(1), 57–67.

International Labour Organization. (1992). *Conditions of work digest: Preventing stress at work*. Geneva: ILO.

Johns, G., & Nicholson, N. (1982). The meanings of absence: New strategies for theory and research. *Research in Organizational Behavior, 4*, 127–172.

Johnson, J. V., & Hall, E. M. (1988). Job strain, work place social support, and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health, 78*, 1336–1342.

Johnson, J. V., Hall, E. M., & Theorell, T. (1989). Combined effects of job strain and social isolation on cardiovascular disease morbidity and mortality in a random sample of the Swedish male working population. *Scandinavian Journal of Work, Environment and Health, 15*, 271–279.

Karasek, R., & Theorell, T. (1990). *Healthy work: Stress, productivity and the reconstruction of working life*. New York, NY: Basic Books.

Kawakami, N., Araki, A., & Kawahima, M. (1997). Effects of work-related stress reduction on depressive symptoms among Japanese blue-collar workers. *Scandinavian Journal of Work, Environment and Health, 23*, 54–59.

Kivimäki, M., Virtanen, M., Vartia, M., Elovainio, M., Vahtera, J., & Keltikangas-Järvinen, L. (2003). Workplace bullying and the risk of cardiovascular disease and depression. *Occupational and Environmental Medicine, 60*(10), 779–783.

Kompier, M. (2003). Job design and well-being. In M. J. Schabracq, J. A. M. Winnubst, & C. L. Cooper (Eds.), *The handbook of work and health psychology* (2nd ed., pp. 429–454). Chichester, UK: Wiley & Sons Ltd.

LaMotagne, A., Keegel, T., Louie, A., Ostry, A., & Landsbergis, P. (2007). A systematic review of the job stress intervention literature 1990–2005. *International Journal of Occupational and Environmental Health, 13*, 268–280.

Landsberger, H. A. (1958). *Hawthorne revisited*. Ithaca, NY: Cornell University Press.

Landsbergis, P. A. (2009). Interventions to reduce job stress and improve work organization and worker health. In P. L. Schnall, M. Dobson, & E. Rosskam (Eds.), *Unhealthy work: Causes, consequences and cures* (pp. 193–209). Amityville, NY: Baywood Publishing.

Landsbergis, P. A., Cahill, J., & Schnall, P. (1999). The impact of "lean" production and related new systems of worker organisation on worker health. *Journal of Occupational Health Psychology, 4*(2), 108–130.

Levi, L. (2000). Legislation to protect worker CV health in Europe. In P. Schnall, K. Belkic, P. Landsbergis, & D. Baker (Eds.), *The workplace and cardiovascular disease* (pp. 269–273). Philadelphia, PA: Hanley & Belfus.

Lim, S., & Cortina, L. M. (2005). Interpersonal mistreatment in the workplace: The interface and impact of general incivility and sexual harassment. *Journal of Applied Psychology, 90*(3), 483–496.

Lindley, R. M. (1996). The school-to-work transition in the United Kingdom. *International Labour Review, 135*(2), 159–170.

Loretto, W., Duncan, C., & White, P. (2000). Ageism and employment: Controversies, ambiguities and younger people's perceptions. *Ageing and Society, 20*(3), 279–302.

Madaus, J. W. (2008). Employment self-disclosure rates and rationales of university graduates with learning disabilities. *Journal of Learning Disabilities, 41*(4), 291–299.

Madaus, J. W., Foley, T. E., McGuire, J. M., & Ruban, L. (2002). Employment self-disclosure of postsecondary graduates with learning disabilities: Rates and rationales. *Journal of Learning Disabilities, 35*, 364–369.

Marchand, A., Demers, A., & Durand, P. (2005). Do occupation and work conditions really matter? A longitudinal analysis of psychological distress experiences among Canadian workers. *Sociology of Health and Illness, 27*, 602–627.

Marine, A., Ruotsalainen, J., Serra, C., & Verbeek, J. (2006). Preventing occupational stress in healthcare workers. *Cochrane Database System Review, 18*(4), CD002892.

Marmot, A. F., Eley, J., Stafford, M., Stansfeld, S. A., Warwick, E., & Marmot, M. G. (2006). Building health: An epidemiological study of "sick building syndrome" in the Whitehall II study. *Occupational and Environmental Medicine, 63*(4), 283–289.

Marmot, M., Feeney, A., Shipley, M., North, F., & Syme, S. L. (1995). Sickness absence as a measure of health status and functioning: From the UK Whitehall II study. *Journal of Epidemiology and Community Health, 49*(2), 124–130.

Martin, A., Sanderson, K., & Cocker, F. (2009). Meta-analysis of the effects of health promotion intervention in the workplace on depression and anxiety symptoms. *Scandinavian Journal of Work and Environmental Health, 35*(1), 7–18.

Mayo, E. (1949). *Hawthorne and the Western Electric Company: The social problems of an industrial civilisation*. London: Routledge.

Mays, V. M., & Cochran, S. D. (2001). Mental health correlates of perceived discrimination among lesbian, gay, and bisexual adults in the United States. *American Journal of Public Health, 91*(11), 1869–1876.

Mays, V. M., Coleman, L. M., & Jackson, J. S. (1996). Perceived race-based discrimination, employment status and job stress in a national sample of Black women: Implications for health outcomes. *Journal of Occupational Health Psychology, 1*, 319–329.

McDermott, E. (2006). Surviving in dangerous places: Lesbian identity performances in the workplace, social class and psychological health. *Feminism and Psychology, 16*(2), 193–211.

McGregor, J., & Gray, L. (2001). *The mature job seeker*. Palmerston North, New Zealand: Massey University.

McKenzie, K. (2003). Racism and health. *British Medical Journal, 326*, 65–66.

Miller, G. V. F., & Travers, C. J. (2005). Ethnicity and the experience of work: Job stress and satisfaction of minority ethnic teachers in the UK. *International Review of Psychiatry, 17*, 317–327.

Morrow, P. C., McElroy, J. C., & Phillips, C. M. (1994). Sexual harassment behaviors and work related perceptions and attitudes. *Journal of Vocational Behavior, 45*, 295–309.

Murray, L. R. (2003). Sick and tired of being sick and tired: Scientific evidence, methods, and research implications for racial and ethnic disparities in occupational health. *American Journal of Public Health, 93*, 221–226.

Nawyn, S. J., Richman, J. A., Rospenda, K. M., & Tonda, L. H. (2000). Sexual identity and alcohol-related outcomes: Contributions of workplace harassment. *Journal of Substance Abuse, 11*(3), 289–304.

Nazroo, J. Y. (1997). *Ethnicity and mental health: Findings from a National Community Survey*. London: Policy Studies Institute.

Nazroo, J. Y. (2003). The structuring of ethnic inequalities in health: Economic position, racial discrimination and racism. *American Journal of Public Health, 93*, 277–284.

Nordenmark, M. (2002). Multiple social roles—Resource or a burden: Is it possible for men and women to combine paid work with family life in a satisfactory way? *Gender, Work and Organization, 9*(2), 125–145.

Norris, W. P. (1991). Liberal attitudes and homophobic acts: The paradoxes of homosexual experience in a liberal institution. *Journal of Homosexuality, 22*, 81–120.

O'Campo, P., Eaton, W., & Muntaner, C. (2004). Labor market experience, work organization, gender inequalities and health status: Results from a prospective analysis of U.S. employed women. *Social Science and Medicine, 58*(2), 585–594.

Palmer, A. (1993). *Some are more equal than others*. London: Stonewall.

Palmore, E. (1990). *Ageism: Negative or positive?* New York, NY: Springer.

Park, K. O., Schaffer, B. S., & Griffin-Blake, C. S. (2004). Effectiveness of a healthy work organization intervention: Ethnic group differences. *Journal of Occupational and Environmental Medicine, 46*, 623–634.

Parker, S. K., Chmiel, N., & Wall, T. D. (1997). Work characteristics and employee well-being within a context of strategic downsizing. *Journal of Occupational Health Psychology, 2*(4), 289–303.

Parker, S. K., Jackson, P. R., Sprigg, C. A., & Whybrow, A. C. (1998). *Organisational interventions to reduce the impact of poor work design*. Sheffield: University of Sheffield/Health & Safety Executive.

Parkes, K. R., & Sparkes, T. J. (1998). *Organisational interventions to reduce work stress: Are they effective? A review of the literature*. Sudbury, ON: HSE Books.

Parks, K. M., & Steelman, L. A. (2008). Organizational wellness programs: A meta-analysis. *Journal of Occupational Health Psychology, 13*(1), 58–68.

Poulsen, K., Jensen, S., Bach, E., & Schostak, J. (2007). Using action research to improve health and the work environment for 3500 municipal bus drivers. *Educational Action Research, 15*(1), 75–106.

Ragins, B. R., & Scandura, T. A. (1995). Antecedents and work-related correlates of sexual harassment: An empirical investigation of competing hypotheses. *Sex Roles, 32*, 429–455.

Raver, C. C. (2003). Does work pay psychologically as well as economically? The role of employment in predicting depressive symptoms and parenting among low-income families. *Child Development, 74*(6), 1720–1736.

Raw, G. J., Roys, M. S., & Leaman, A. (1993, August). Sick building syndrome, productivity and control. *Property Journal,* 17–19.

Romeri, E., Baker, A., & Griffiths, C. (2007). Alcohol-related deaths by occupation, England and Wales, 2001–5, Office for National Statistics. *Health Statistics Quarterly, 35*. Retrieved from http://www.statistics.gov.uk/downloads/theme_health/HSQ35.pdf

Rosenheck, R., Leslie, D., Keefe, R., McEvoy, J., Swartz, M., & Perkins, D. (2006). Barriers to employment for people with schizophrenia. *American Journal of Psychiatry, 163*, 411–417.

Rospenda, K. M., & Richman, J. A. (2005). Harassment and discrimination. In J. Barling, E. K. Kelloway, & M. R. Frone (Eds.), *Handbook of work stress* (pp. 149–188). Thousand Oaks, CA: Sage.

Rospenda, K. M., Richman, J. A., & Shannon, C. A. (2009). Prevalence and mental health correlates of harassment and discrimination in the workplace: Results from a national study. *Journal of Interpersonal Violence, 24*(5), 819–843.

Rospenda, K. M., Richman, J. A., Wislar, J. S., & Flaherty, J. A. (2000). Chronicity of sexual harassment and generalized work-place abuse: Effects on drinking outcomes. *Addiction, 95*, 1805–1820.

Rugulies, R., Bültmann, U., Aust, B., & Burr, H. (2006). Psychosocial work environment and incidence of severe depressive symptoms: Prospective findings from a 5-year follow-up of the Danish work environment cohort study. *American Journal of Epidemiology, 163*(10), 877–887.

Saks, A., & Waldman, D. (1998). The relationship between age and job performance evaluations for entry-level professionals. *Journal of Organizational Behavior, 19*(4), 409–419.

Sanderson, K., & Andrews, G. (2006). Common mental disorders in the workforce: Recent findings from descriptive and social epidemiology. *Canadian Journal of Psychiatry, 51*(2), 63–75.

Sargeant, M. (2001). Lifelong learning and age discrimination in employment. *Education and the Law, 13*(2), 141–154.

Saunders, T., Driskell, J. E., Johnston, J. H., & Salas, E. (1996). The effect of stress inoculation training on anxiety and performance. *Journal of Occupational Health and Psychology, 1*(2), 170–186.

Schnall, P. L., Pieper, C., Schwartz, J. E., Karasek, R. A., Schlussel, Y., Devereux, R. B., et al. (1990). The relationship between "job strain", workplace diastolic blood pressure, and left ventricular mass index: Results of a case-control study. *JAMA, 263*(14), 1929–1935.

Schulze, B., & Angermeyer, M. C. (2003). Subjective experiences of stigma: A focus group study of schizophrenia patients, their relatives and mental health professionals. *Social Science & Medicine, 56*, 299–312.

Schuman, E., & Kleiner, B. (2001). Is age a handicap in finding employment? *Equal Opportunities International, 20*(5–7), 48–52.

Semmer, N. (2006). Job stress interventions and the organisation of work. *Scandinavian Journal of Work, Environment and Health, 32*(6), 515–527.

Siegrist, J., Peter, R., & Junge, A. (1990). Low status control, high effort at work and ischemic heart disease: Prospective evidence from blue collar men. *Social Science and Medicine, 31*, 1127–1134.

Skidmore, P. (2004). A legal perspective on sexuality and organization: A lesbian and gay case study. *Gender, Work and Organization, 11*(3), 229–253.

Smiley, E. (2005). Epidemiology of mental health problems in adults with learning disability: An update. *Advances in Psychiatric Treatment, 11*, 214–222.

Smith, A., Johal, S., Wadsworth, E., Davey, S. G., & Peters, T. (2000). *The scale and impact of occupational stress: The Bristol Stress and Health at Work Study*. London: Health & Safety Executive Research Report No. 265.

Snape, E., & Redman, T. (2003). Too old or too young? The impact of perceived age discrimination. *Human Resource Management Journal, 13*(1), 78–89.

Sorensen, L., & Roberts, S. J. (1997). Lesbian uses of and satisfaction with mental health services: Results from Boston Lesbian Health Project. *Journal of Homosexuality, 33*(1), 35–49.

Sorensen, G., Stoddard, A., Ockene, J. K., Hunt, M. K., & Youngstrom, R. (1996). Worker participation in an integrated health promotion/health protection program: Results from the WellWorks Project. *Health Education Quarterly, 23*, 191–203.

Stansfeld, S. A., North, F. M., White, I., & Marmot, M. G. (1995). Work characteristics and psychiatric disorder in civil servants in London. *Journal of Epidemiology and Community Health, 49*(1), 48–53.

Stansfeld, S. A., Bosma, H., Hemingway, H., & Marmot, M. G. (1998). Psychosocial work characteristics and social support as predictors of SF-36 health functioning: The Whitehall II study. *Psychosomatic Medicine, 60*, 247–255.

Stansfeld, S. A., & Candy, B. (2006). Psychosocial work environment and mental health: A meta-analytic review. *Scandinavian Journal of Work and Environmental Health, 36*(6), 443–462.

Stansfeld, S. A., Clark, C., Caldwell, T., Rodgers, B., & Power, C. (2008). Psychosocial work characteristics and anxiety and depressive disorders in midlife: The effects of prior psychological distress. *Occupational and Environmental Medicine, 65*(9), 634–642.

Stansfeld, S. A., Rasul, F. R., Head, J., & Singleton, N. (2011). Occupation and mental health in a national UK survey. *Social Psychiatry and Psychiatric Epidemiology, 46*(2), 101–110.

Stansfeld, S. A., Woodley Jones, D., Rasul, F., Head, J., Clarke, S., & Mackay, C. (2008). Work related distress in the 1990s: A real increase in ill health? *Journal of Public Mental Health, 7*, 26–31.

Stuart, H. (2004). Stigma and work. *Healthcare Papers, 5*, 100–111.

Stuart, H. (2006). Mental illness and employment discrimination. *Current Opinion in Psychiatry, 19*(5), 522–526.

Tucker, P., & Knowles, S. R. (2008). Review of studies that have used the Standard Shiftwork Index: Evidence for the underlying model of shiftwork and health. *Applied Ergonomics, 39*(5), 550–564.

van der Klink, J. J., Blonk, R. W., Schene, A. H., & van Dijk, F. J. (2001). The benefits of interventions for work-related stress. *American Journal of Public Health, 91*, 270–276.

van Vegchel, N., de Jonge, J., Bosma, H., & Schaufeli, W. (2005). Reviewing the effort-reward imbalance model: Drawing up the balance of 45 empirical studies. *Social Science and Medicine, 60*(5), 1117–1131.

Velligan, D. I., & Gonzalez, J. M. (2007). Rehabilitation and recovery in schizophrenia. *Psychiatric Clinics of North America, 30*(3), 535–548.

Virtanen, M., Pentti, J., Vahtera, J., Ferrie, J. E., Stansfeld, S. A., Helenius, H., et al. (2008). Overcrowding in hospital wards as a predictor of antidepressant treatment among hospital staff. *American Journal of Psychiatry, 165*(11), 1482–1486.

Wadsworth, E., Dhillon, K., Shaw, C., Bhui, K., Stansfeld, S., & Smith, A. (2007). Racial discrimination, ethnicity and work stress. *Occupational Medicine (London), 57*(1), 18–24.

Walker, A. (2005). Towards and international political economy of ageing. *Ageing and Society, 25*, 815–839.

Wieclaw, J., Agerbo, E., Mortensen, P. B., & Bonde, J. P. (2006). Risk of affective and stress related disorders among employees in human service professions. *Occupational and Environmental Medicine, 63*(5), 314–519.

Wieclaw, J., Agerbo, E., Mortensen, P. B., Burr, H., Tuchsen, F., & Bonde, J. P. (2008). Psychosocial working conditions and the risk of depression and anxiety disorders in the Danish workforce. *BMC Public Health, 7*(8), 280.

Wing, J. K. (1970). Unemployment among chronic psychotic patients in Camberwell: Their need for rehabilitation. *Proceedings of the Royal Society of Medicine, 63*(12), 1329–1332.

Wood, G., Wilkinson, A., & Harcourt, M. (2008). Age discrimination and working life: Perspectives and contestations—A review of the contemporary literature. *International Journal of Management Reviews, 10*(4), 425–442.

Woods, J. D. (1993). *The corporate closet: The professional lives of gay men in America*. New York, NY: The Free Press.

Ybema, J. F., & van den Bos, K. (2010). Effects of organizational justice on depressive symptoms and sickness absence: A longitudinal perspective. *Social Science and Medicine, 70*(10), 1609–1617.

Ylipaavalniemi, J., Kivimäki, M., Elovainio, M., Virtanen, M., Keltikangas-Järvinen, L., & Vahtera, J. (2005). Psychosocial work characteristics and incidence of newly diagnosed depression: A prospective cohort study of three different models. *Social Science and Medicine, 61*(1), 111–122.

Chapter 11
An Integrated Prevention Science Model: A Conceptual Foundation for Prevention Research

Hanno Petras and Zili Sloboda

Introduction

Prevention science is transdisciplinary that includes the broad domains of epidemiology, intervention development and implementation, and research design and statistical methods (see Fig. 11.1). Prevention science is characterized by complex feedback loops between research and practice (Coie et al., 1993). The objective of prevention science is to test theory-based hypotheses toward enhancing or refining the process that makes individuals' engagement in healthy and non-risky behaviors more likely (Biglan et al., 2011). These theories explain the mechanisms of the prevention process related to specifying intervention targets, identifying preventive activities to address the targets, and determining the most appropriate design, measurements, and analyses to evaluate the process.

Importantly, intervention efforts are commonly challenged by individual and group differences across both time and place affecting readiness to respond to and benefit from the intervention. These differences are the consequence of individual vulnerability for negative and unhealthy behaviors. Vulnerability may be the result of failure to achieve appropriate developmental benchmarks that impede healthy decision-making in stressful or risky situations. As such, one may be at risk at any time over the life course. In addition to differences in the developmental stages with respect to vulnerability, it is also widely acknowledged that individuals are embedded in more proximal and distal ecological domains (e.g., the family, school, peer network, community), which indicate that the success of prevention efforts is also moderated by environmental features. Finally, prevention views individuals from a

The authors are grateful for the thoughtful comments provided by Moshe Israelashvili on an earlier version of this chapter.

H. Petras (✉)
JBS International, Inc., 5515 Security Lane, Suite 800, North Bethesda, MD 20852-5007, USA
e-mail: hpetras@jbsinternational.com

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_11,

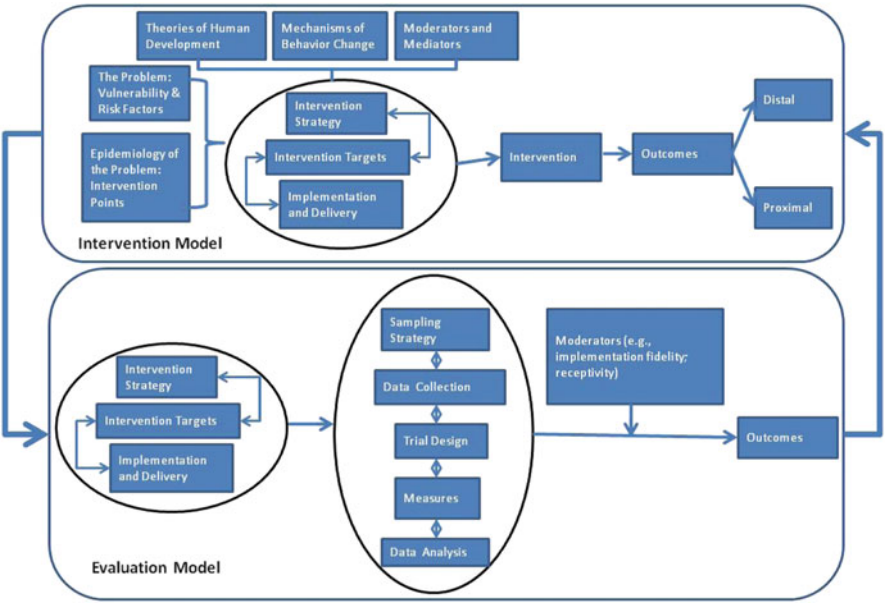

Fig. 11.1 Prevention science cycle

human agency perspective, that is, individuals are not just influenced by environmental toxins, but importantly interact with the environment based on their motivation, intentions, and self-efficacy.

Prevention Science Past and Future

The etiological and preventive knowledge that has been accumulated over the last 30 years was made possible through parallel and cross-cutting research that has striven to determine the causal pathways to mental, emotional, and behavioral (MEB) disorders. As prevention is central to a public health framework, it has guided early prevention research efforts. This framework, based on an infectious disease model, arrays interventions that address the risk statuses within a population—those who are at risk for or susceptible to infection (primary), those who experience the early signs and symptoms of infection (secondary), and those who already have the infection (tertiary). Within this framework, prevention efforts focus on the three legs of prevention—the host, the agent, and the environment (Haddon, 1972). Prevention efforts then would "immunize" or "inoculate" the host or susceptible groups, eliminate the source of the infection (virus or bacteria), or implement environmental modification to make the physical and social environment less suitable for the agent. Indeed, early prevention efforts, particularly those designed for substance use prevention, attempted to "inoculate" young adolescents against the initiation of substance use, particularly tobacco use (Evans et al., 1978).

The longitudinal research required to provide insights into risk factors associated with many MEB disorders was not available until the early to mid-1970s. The findings from this research resulted in two major research efforts that, although related, have different implications. The first effort relies heavily on risk and indicators of risk (i.e., determining the characteristics that differentiate those who go on to having an MEB disorder from those who do not), in many instances controlling for age, race/ethnicity, gender, and other environmental conditions (Hawkins, Catalano, & Miller, 1992). The other train of research examines populations that are considered vulnerable for a variety of reasons or circumstances such as loss of one or two parents or family members, war, natural disasters, and/or other environmental disorganization (Werner, 1989).

It should be noted that none of these groups negates the importance of the other, but each suggests that its perspective has different implications for prevention. The first has dominated and inspired the Institute of Medicine (IOM) (Mrazek & Haggerty, 1994) to advance a new risk-based rubric for prevention programming. This rubric includes universal prevention strategies that target full populations and that evidence some benefits for all outweighing any iatrogenic risks of negative consequences; selective prevention strategies that target subpopulations or groups that are potentially at elevated risk for a disorder but that do not exhibit signs or behaviors associated with the disorder; and, finally, indicated prevention strategies that target those subpopulations that are assessed at elevated risk for a disorder but that do not evidence any associated symptoms of the disorder. The second approach has had less traction in the field but has advanced programs that enhance protective or resiliency particularly among those most at risk. These programs have been shown to have an impact, particularly in life skills development that promotes bonding to family and prosocial groups and to social, emotional, cognitive, behavioral, and moral competence; foster self-determination; and so forth (see Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2004; Catalano, Hawkins, Berglund, Pollard, & Arthur, 2002).

Two landmark publications by Hawkins and colleagues (1992) and Glantz and Pickens (1992) are representative of the risk factor approach. Hawkins and colleagues (1992) reviewed and summarized risk factors "evidence" from both epidemiologic and intervention research studies reported in the literature. They divided these risk factors into two categories: those that are related to society and culture or contextual factors and those that pertain to the individual and his/her intrapersonal environment. The contextual factors include laws and norms favorable to substance use behaviors, factors related to availability, those related to extreme economic deprivation, and those related to neighborhood disorganization. The individual and intrapersonal factors include physiological issues, family history of substance use and attitudes toward substance use, poor/inconsistent family management, family conflict, low family bonding, early/persistent problem behaviors, academic failure, low commitment to school, peer rejection during elementary grades, association with substance-using peers, alienation and rebelliousness, and early onset of drug (substance) use. In this same article, the concept of protective factors associated with substance use is presented and includes such

items as parental attachment and conventionality (Brook, Brook, Gordon, Whiteman, & Cohen, 1990; Hawkins et al., 1992), although the authors point out that there was a dearth of research on protective factors at that time. While the work of Hawkins and colleagues addresses substance use and to some extent substance abuse, Glantz and Pickens (1992) were more concerned about factors related to substance abuse or substance use disorders with a focus on intrapersonal or biological, psychological, and psychiatric mechanisms.

Interest in protective factors that serve to differentiate those who share similar vulnerabilities but who experience differing MEB disorder trajectories is well represented in the works of researchers such as Garmezy (1985), Garmezy (1991), Masten and colleagues (1990), Redmond et al. (2009), Rutter (1985), Rutter et al., (1997), and Werner (1989) who examined protection or resiliency. These researchers define resiliency as a process, capacity, or successful adaption within difficult, challenging, or threatening situations and resiliency factors as those characteristics or processes that play a role in mediating or moderating the effects of exposure to risk.

As important as the concept of risk and protective factors has been to the progress made in prevention programming, few attempted to determining if these factors group into meaningful clusters (Sloboda, Glantz, & Tarter, 2012) or whether they elucidate the actual relationship to MEB disorders per se (Catalano, Kosterman, Hawkins, Newcomb, & Abbott, 1996; Lacourse et al., 2002). We do know that these risk factors likely represent indicators of other underlying mechanisms that have not been sufficiently explored. For instance, those risk factors that represent individual or intrapersonal risk factors mostly capture constitutional (biopsychological) characteristics such as low IQ, hyperactivity, and concentration problems that likely have developmental and neurobiological origins (Glantz & Pickens, 1992; Hawkins et al., 1992). In addition, behavioral and attitudinal elements such as antisocial behavior, beliefs, and attitudes regarding drug use are also represented as individual/peer risk factors. These are independent factors within the Hawkins individual/peer domain, but the bidirectional influences are obvious. The other risk factor domains (family, school, and community) capture aspects that impede positive socialization experiences due to challenged socialization agents or intrapersonal mechanisms that hinder socialization, such as family conflict, academic failure, failure to attach and commit to prosocial organizations, and low neighborhood attachment, that act as stressors on individual and family performance. The concepts of risk and protection and resiliency suggest the notion of vulnerability, being susceptible or predisposed, or being at risk (i.e., making poor decisions about the level of risk for engagement in a behavior that may have negative social or health outcomes). Furthermore, the findings from this research across MEB disorders generally indicate that, unlike most infectious diseases, these disorders have multiple precursors that often follow different paths or trajectories across the lifespan, thus differing dramatically in outcomes and complexity (Herrenkohl et al., 2000; Jessor & Jessor, 1977).

The IOM 2009 report (O'Connell, Boat, & Warner, 2009) refers to the definition of competence of Masten, Burt, and Coatsworth (2006) that considers competence a "family of constructs" that reflect a capacity for achieving age-salient effective

adaptive outcomes within the individuals' cultural and historical context. Kellam and colleagues (1975) lay these competencies out along a life-course perspective showing the adaptive tasks required within different social fields including the family, school, peers, community, and work settings. The achievement in competencies early in the life course serves to improve the achievement of competencies later in life.

Although the concept of vulnerability has gained more attention in recent years (O'Connell et al., 2009), the field has not developed a standard definition or set of measures for the concept (Cicchetti & Blender, 2006). In general, the concept encompasses the attainment of age-related competencies. More attention is being given to the development of benchmarks for these competencies (see chapters in this book by Graber, Hill, & Saczawa, 2013; Kim-Spoon & Farley, 2013; Rauh & Bergmann, 2013; Rebok, Parisi, & Kueider, 2013).

Etiologic Theories and Socialization

The fact that many risk and unhealthy behaviors are associated with multiple factors has led to the development and application of a variety of etiologic theories, which are often complementary, but nevertheless exist in parallel universes. Depending on the researcher's discipline (e.g., sociology, psychology, biology), different pieces of the puzzle are investigated using different constructs, measurements, samples, and analytic methods. For example, Petraitis, Flay, and Miller (1995) identified 14 different multivariate theories explaining early stage substance use emphasizing cognitions, social learning processes, commitment to prosocial values, and institutions as well as intrapersonal processes. In concert with a lack of comparison and integration of these existing theories, the guidance for the analysis of etiological data and for the design of preventive interventions has been challenged and their international applicability is unknown (Catalano et al., 2012).

To this end, Oetting and colleagues (Oetting, Deffenbacher, & Donnermeyer, 1998; Oetting & Donnermeyer, 1998; Oetting, Donnermeyer, & Deffenbacher, 1998; Oetting, Donnermeyer, Trimble, & Beauvais, 1998) proposed primary socialization theory as an integrated theory for prevention science. There are many definitions of socialization available. The common theme to all of them is that socialization represents a social and psychological interactive process between the individual and primary (and secondary) socialization sources. The focus of this process is on the internalization of societal goals, normative behaviors and cues, values associated with these behaviors, and roles and responsibilities as members of a society. In any society, certain primary socialization agents are responsible for the individual's internalization societal goals, norms, and values. In Western societies and most likely in other societies as well, these agents are usually the family, the school, and the peer clusters (see Fig. 11.2). Recent research has argued that the foundation of these foci is communicated during the early socialization phase through parent–child interactions in five distinct domains, including protection,

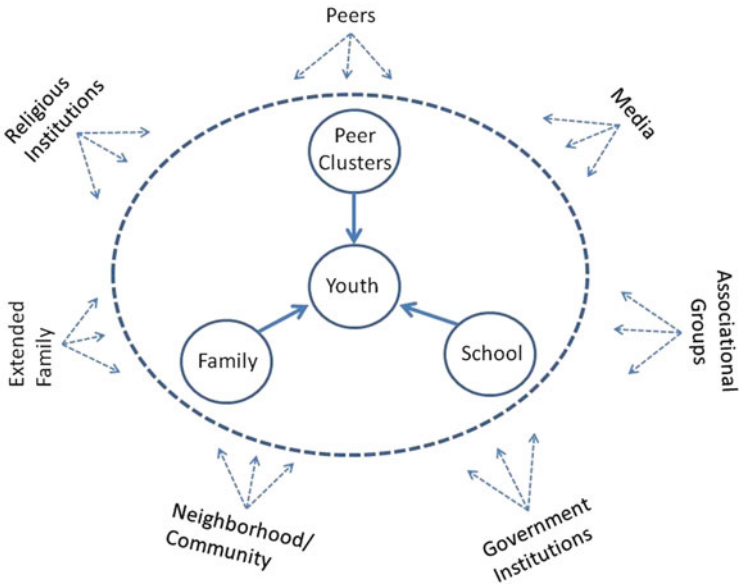

Fig. 11.2 Primary and secondary socialization agents (Leukefeld & Leukefeld, 1999, modified)

mutual reciprocity, control, guided learning, and group participation (Grusec & Davidov, 2010). Internalization of these foci occurs through the provision of positive and negative sanctions depending on whether these requirements are met. It is important to recognize that the socialization process is bidirectional in that individuals are active agents in the process (Kuczynski, 2003). Specifically, individuals differ in their biological preparedness to be socialized and to process relevant information. For example, individuals may have deficiencies in executive functioning and other cognitive impairments, which may make the communication and internalization process more taxing. For example, there is considerable empirical (Jensen, Martin, & Cantwell, 1997; Lahey, McBurnett, & Loeber, 2000; Satterfield & Schell, 1997) and theoretical (e.g., Loeber et al., 1993; Moffitt, 1993; Patterson, DeBaryshe, & Ramsey, 1989) evidence linking ADHD and concentration problems to such socialization challenges. Patterson, Reid, and Dishion (1992) present a good example for this interactive process. They argue that individuals who experience early childhood impairment (i.e., neurological deficits such as attention problems and social problems such as coercive parenting, peer rejection, and school failure) are set on a trajectory characterized by high levels of physical aggression beginning in early childhood and continuing throughout development, resulting in criminal and violent behavior in adolescence and adulthood.

Importantly, weak bonds to socialization agents hinder the communication of societal norms, decrease the individual's commitment to prosocial institutions, and increase the likelihood to interact with deviant peers. All of these processes are recognized risk factors for engaging in deviant and unhealthy behaviors. However,

as suggested by Oetting (1999), risk and protective factors should be organized according to their influence on the primary socialization process and in particular with respect to the extent that they hinder social bonding or prevent the communication of prosocial norms.

The socialization process, the number and types of socialization agents involved, and the roles of the socialization agents themselves change depending on the age and competencies of the individual, with the role changing from one that is directive to one that is more supportive as discussed above. Primary socialization agents generally include the mother or caregiver and the extended family. Secondary socialization agents include school staff and policies, peers, church-related individuals and doctrine, community members and regulations, work-related colleagues and policies, media content, and societal laws. Importantly, while individuals initially (infancy through young adulthood) do not interact with these secondary socialization agents directly, they become over their life course more exposed to and influenced by them.

Socialization and Decision-Making

While some theories view deviant behavior as the result of a rational evaluation process in which costs and benefits of the involved behavior are weighted (Ajzen, 1988; Janz & Becker, 1984; Rogers, 1983), primary socialization theory argues that behaviors, attitudes, and beliefs emerge concurrently from the individual's interaction with the primary and later secondary socialization sources. In other words, while deviant and unhealthy behaviors are learned through active and observational learning, the cognitive representation of the social interaction with socialization agents is the ultimate determinant of individual behavior (Guerra, Nucci, & Huesmann, 1994). Consequently, optimal socialization leads to informed decisions regarding acceptable behaviors when confronted with verbal, behavioral, emotional, and environmental cues.

We have focused on primary socialization given its documented importance for an individual to acquire the fundamental beliefs and knowledge to behave according to a society's culture. While it is recognized that socialization is a lifelong process, also other processes exist that contribute to the individual's attitudes, beliefs, and knowledge with respect to decision-making. In a recent publication, Israelashvili (2012) took on the task to propose a conceptual differentiation between related concepts (i.e., adjustment, adaptation, coping, and socialization). While it is beyond this chapter to discuss these concepts further, the first three concepts address situations in which the individual does not have access to optimal cognitive schemata and a renegotiation process between environmental demands and individual abilities and knowledge is set into motion. This may lead to changes in attitudes and beliefs, increases in the behavioral repertoire, and changes in preferred responses to environmental demands.

Integrating over several behavior change models (i.e., Social Information Processing Theory, Social Learning Theory, and Integrated Behavioral Model), the conceptual underpinning of the decision-making process highlights the strong interaction of biological, psychological, and socio-structural factors that inform how decisions are made and acted on. Major components of the behavioral decision-making process include verbal, behavioral, emotional, and environmental cues and draw on learned repertoires of behaviors. The process itself consists of encoding and evaluating environmental cues and selecting from individually stored behaviors to determine which behavior is the most efficacious to achieve the desired and goal-consistent outcomes (Crick & Dodge, 1994; Lemerise & Arsenio, 2000).

Social Representation of Cues

Social cues are at the core of the decision-making process. Social Information Processing Theory (SIPT; Crick & Dodge, 1994; Dodge & Crick, 1990; Lemerise & Arsenio, 2000) helps explain how social cues prompt behavioral decisions. Social cues may be verbal or behavioral themselves, but cues may also originate from the environment in terms of affordances (Gibson, 1979). These cues may trigger an array of potential behaviors that form a repertoire from which individuals draw those felt to be most appropriate. The selection process is not always spontaneous or reflective and is complex. SIPT hypothesizes that a behavioral response is based on a process involving the above-described activities. As an example, imagine a child is verbally attacked by a classmate and the child needs to classify the event (i.e., attention, encoding) and determine why it happened (i.e., attribution of cause). The child then needs to clarify the goals appropriate for this situation. For example, he/she could maintain a friendly relationship with the classmate or express that the behavior is not tolerated. Finally, the child generates possible responses for the situation and evaluates them in terms of anticipated outcomes, relations to goal(s), and skills or efficacy for enacting the response. The child may consider responding by physically attacking the other student but ultimately will reject this response to prevent the situation from escalating or out of fear that he/she lacks the appropriate physical strength (lack of efficacy) in an assault on the classmate. While earlier SIPT models assumed the process developed sequentially, it is now believed that the different stages of information processing may occur in parallel fashion with the opportunity for several feedback loops.

Given that individuals in complex societies are exposed to a multitude of social cues prompting a behavioral response on a daily basis, most do not execute the complete information processing cycle but rather apply cognitive models of relationships drawing on cues experienced in the past and memories of the outcomes of particular expressed behavioral responses. Familiar situations often prompt quasi-automatic or habitual behavioral responses. While this approach is efficient, it can result in judgment and reasoning errors depending on the generality and quality of the individual's social knowledge and behavioral repertoire. Finally, in addition to

behaviors based on social knowledge as well as active processing, individuals may apply preemptive or script-based processing, which is rapid, automatic, and nonrational in situations that are highly emotionally arousing (e.g., religious or spiritual in nature, interpersonal relationships, matters of great importance or value). Evidence exists to suggest that individuals with low self-regulatory ability apply a preemptive processing style more frequently.

Emotionality and the Decision-Making Process

Standard models of decision-making have typically focused on the effortful, intense-processing, and affectively neutral decision-making process (Chang & Sanfey, 2008) despite the accumulating evidence of the existence of a dual-process model (Poldrack & Packard, 2003). Building on the SIPT model (Crick & Dodge, 1994), Lemerise and Arsenio (2000) extended SIPT to integrate affect and cognition. They argue that individuals differ not only in their level of biological competencies and database of past experience but also in their emotional style and regulatory abilities, both of which influence the processing of social/emotional information and decision-making in stressful situations. Damasio (1994) has argued that affect may increase the efficiency of the information processing and decision-making process by attaching somatic markers to specific behaviors or cognitions. If a person is confronted with a similar situation, these somatic markers will narrow or reduce the array of possible alternatives and thus will lead to increases in efficiency. Importantly, experiences are important for fine-tuning the impact of affect on the processing and decision-making process (Damasio, 1994). However, individual variations in temperament and emotion regulation are important moderators influencing the extent to which this will introduce any biases.

Developmental Aspects of the Decision-Making Process

The quality of the decision-making process depends on the individual's quality of cognitive skills (e.g., experiential knowledge, attentional abilities, and mental organizational skills) and the ability to process information and social cues, all of which evolve developmentally over time.

Children's earliest interactions occur in the family before they reach school. They may encounter risk when they experience interaction with parents or caregivers who fail to nurture and have ineffective parenting skills in a chaotic family setting or the caregiver is a substance abuser or a person with mental illness. Sufficient evidence is available showing that the consequences of mothers' intake of alcohol, nicotine, and drugs during pregnancy negatively affect developing fetuses. Such deficiencies impede reaching significant developmental competencies and make a child vulnerable and at risk for negative behaviors later on. By age 2 or

3, children begin manifesting disruptive behaviors and temper tantrums, are disobedient, or demonstrate destructive behaviors. If not properly addressed, these personality traits and behaviors can become problematic later in life.

During middle childhood, increasingly more time is spent away from the family most often in school and with same-age peers. Therefore, competencies developed during infancy and early childhood will greatly influence the extent to which the school-age child will cope and bond with school and prosocial attitudes and their involvement in prosocial activities. Mental disorders that have their onset during this period, such as anxiety disorders, impulse control disorder, and conduct disorders, may also impede the development of healthy attachment to school, cooperative play with peers, adaptive learning, and self-regulation. Children of dysfunctional families often affiliate at this time with deviant peers, thus putting themselves at risk for negative life choices, including drug use, alcohol misuse, and involvement in illegal activities.

Adolescence is a developmental period when the youth are exposed to new ideas and behaviors through increased associations with people and organizations beyond those experienced in childhood and is a time to "try out" adult roles and responsibilities. The desire to assume adult roles and more independence at a time when significant changes are occurring in the adolescent brain also creates a potentially opportune time for poorly thought-out decisions and involvement in potentially harmful behaviors, such as risky sexual behaviors, smoking and drinking, risky driving behaviors, and illicit drug use.

Determinants of the Decision-Making Process

Asserting that behavioral intentions are the best determinants for behavior, the Theory of Reasoned Action (TRA; Ajzen & Fishbein, 1980), the Theory of Planned Behavior (TPB; Ajzen, 1991), and the Integrated Behavioral Model (IBM; Fishbein, 2000, 2008) argue that intentions are influenced by attitudes toward performing a behavior, norms associated with the behavior, and perceived control over the performance of the behavior considering other situational factors. The individual's attitudes toward a particular behavior are influenced by the attributed consequences of that behavior, which are weighted by positive or negative evaluations of that outcome. Normative beliefs or perceptions of the frequency by which parents, friends, and peers are likely to perform the particular behavior tap into perceptions of social normative pressures in favor or against the behavior. Returning to the exemplar child mentioned above who had been attacked by a classmate, norms and beliefs are hypothesized to influence every step of the decision-making process. The likelihood of the individual responding with violence will depend on his or her evaluation of consequences for such a response, the standards for conflict resolution in his/her context, the extent to which teachers, peers, and parents will approve of a particular response, and the extent to which such approval matters to the individual. The last concept that will influence the

intentions to perform the behavior as well as the behavior itself is represented by the individual's perceived control over the performance of the behavior. This concept is similar to Bandura's (1997) self-efficacy construct that includes perceived amount of control over and confidence in the performance of a behavior.

Complementary to TRA and TRB, Social Learning Theories (SLT; Bandura, 1977) explain human behavior as a reciprocal interaction among cognitive, behavioral, and environmental determinants. SLT emphasizes that positive behavior is modeled by observing and imitating the behavior of significant others in a particular context. When displaying newly acquired behavioral skills, individuals experience positive support from the environment. This support increases perceived capability and confidence and facilitates positive attitudes toward implementing these and other related new skills.

In summary, TRB and TRA formulate hypotheses about the effects of proximal determinants related to decision-making, behavioral intentions, and behavior. SLT, on the other hand, focuses on the origins of attitudes, norms, and skills, which are the result of the context-specific cognitive representation of the individual–environment interaction.

Vulnerability and Decision-Making

Recent advances in neuroscience have documented the high level of neuroplasticity or reorganization that occurs within and between brain structures regulating emotional (limbic system) and cognitive functioning (prefrontal cortex) from early childhood through late adolescence into young adulthood. As children grow older, their amount of social experiences is likely to increase in both a quantitative (e.g., acquiring new strategies to deal with social situations) and a qualitative (e.g., learning more skillful and adaptive ways for negotiating situations) fashion. This increase in social experience is likely due to the increases in the processing of social stimuli, increased exposure to new situations for which a response template is not available in the long-term memory, and continuing socialization efforts by a large array of agents regarding appropriate behavior with varying outcomes that may or may not be shaped by each occurring situation. Studies have shown that as children grow older they improve their accuracy in assessing social cues and their efficiency in responding to them, from very simple cues to those that are more nuanced and complex. These enhancements in processing skills are complimented by developing more effective ways of representing and organizing social information the short- and long-term memory. Improvements in the organization of cues will be more salient for those social contexts with which individuals are more familiar. In addition to increases in processing efficiency, there is evidence that, as individuals grow older, processing patterns become more rigid with early socialization experiences being more influential in later decision-making.

On the other hand, vulnerability during adolescence appears to be very prominent for making poor decisions because during this period the body and brain are

maturing differentially. It is at this time that the brain's prefrontal cortex undergoes enormous changes. The prefrontal cortex is responsible for executive cognitive functions (e.g., decision-making, self-monitoring, abstract thinking, and forming goal strategies) and oversees behavioral and affective regulation. It is also the last region of the brain to mature functionally. Casey and colleagues (2010) suggest that this is not the only process of change at this time, pointing out that, if that were the case, adolescents would not be different from their younger peers, particularly emotionally. These researchers feel that there is an imbalance in brain development resulting in a disparity in brain systems responsible for affective processing or emotions: the cortical regions of the brain (Fareri, Martin, & Delgado, 2008). These changes along with hormonal and other normal biological changes at the same time present additional challenges to adolescents. The variation in development across same-age peers leads to differential expectations with those appearing more mature to be viewed as more capable of assuming adult roles and those whose physical appearance has not evidenced maturity to be viewed as not being able to assume similar roles. In summary, then, uneven brain system development establishes stresses and erratic emotions that along with variability in physical development provide opportunities for poor decision-making and increase the engagement in risky behaviors.

Interest in the propensity to risk taking focused on the idea that some individuals were predisposed to engaging in behaviors that had the potential for negative outcomes. This led to research that identified the personality characteristics of people who were more often involved in risky behaviors. The most influential work in this area involves the sensation-seeking personality construct described by Zuckerman (1994) as "a trait defined by the seeking of varied, novel, complex, and intense sensations and experiences and the willingness to take physical, social, legal and financial risks for the sake of such experiences" (p. 27). Many studies have since been conducted to determine the role of high sensation seeking in risky behavior, including the use of alcohol, tobacco, and other drugs (Romer & Hennessy, 2007); high-risk sexual behavior (Romer et al., 2009); and risky driving behavior (Jonah, 1997). Zuckerman (1994) places the origin of sensation seeking in the dopamine brain pathways that drive interest in novel and rewarding activities. Generally, this brain activity increases during adolescence, when it peaks, then gets reduced in young adulthood with general maturity, experience, and recognition of the potential consequences of risk-taking behavior (Romer et al., 2009).

Steinberg (2010) suggests that reward seeking and impulsivity develop differentially and are related to different neural systems. Furthermore, as indicated above, the timing of these system evolutions also is quite different with reward seeking increasing between pre- and mid-adolescence and decreasing in late adolescence into adulthood. However, impulsivity tends to decline from age 10 on. Consequently Steinberg postulates "risk taking in middle adolescence may be due to the combination of relatively higher inclinations to seek rewards and still maturing capacities for self-control" (Steinberg, 2010, p. 216).

In addition to variation in the process of maturation, recent advances in brain imaging have helped scientists understand more on how brain injuries or

impairments can lead to risk-taking behaviors. With the new technologies, neuroscientists have identified the neural circuitry involved in decision-making and reward-related processing within the frontal corticobasal ganglia networks. Fareri and colleagues (2008) review the research to date that supports this finding and the development of these circuits during adolescence.

To summarize, vulnerability with respect to decision-making addresses an array of interactive factors, including aspects of normal brain development and maturation, variability in personality traits and reward-seeking behavior, and structural and functional brain impairments. These diverse sources of influence have in common that they influence certain aspects of or the entire information processing and decision-making cycle by increasing the likelihood of biased information processing and storage and rapid and nonrational goal and behavior selection.

Prevention as a Socialization Agent or Enhancer

Key to the integrated prevention science model is the socialization process that serves as an important foundation for individuals' social information processing and decision-making. As societies become more complex and dynamic, individuals experience diverse networks of socialization agents, which increase the likelihood that they will be exposed to potentially mixed sets of expectations regarding behaviors and cues. It is in these situations that prevention takes on two roles: to reinforce the role of other socialization agents or to assume the role of socialization agent per se.

Objectives and Processes of Prevention (Intervention Targets and Implementation Issues)

Engagement in prosocial behaviors is dependent on a balance between embracing societal- and contextual-specific expectations and having the competencies to do so. Prevention interventions, like socialization agents, have as their goal the internalization of acceptable behaviors that are appropriate for the general society and for specific contexts of everyday life. As discussed above, this process includes modeling and trying on behaviors within a lifeworld of experiences and belief and value systems, understanding social cues, and having the competencies and skills to successfully performing the behaviors.

The major difference between socialization agents and prevention interventions is that the latter formalize and structure the essential elements of the socialization process following evidence-based principles. Prevention interventions are viewed then as preparing individuals by establishing or reinforcing prosocial attitudes, beliefs, and goals. In addition, these interventions provide skills for the appropriate

response to challenging situations based on the interpretation and evaluation of situational cues such as the availability of tobacco, alcohol, and marijuana or opportunities to shoplift or perform poorly in school to enhance making appropriate and healthy decisions. These aspects of prevention intervention are informed by a variety of available theoretical perspectives discussed above representing social control as well as social learning, the fundamental mechanisms of socialization.

Although the ultimate aim of prevention is the performance of positive or non-risky behaviors, the outcomes of prevention focus mainly on the prevention process, behavioral intent, and the acquisition of skills to perform the appropriate behavior with efficacy. The intentions are the result of the decision-making process: gathering, interpreting, and evaluating information within the normative context that supports sets of alternative behaviors.

Prevention targets individuals either directly through a formal intervention program or through existing socialization agents. Examples of the former include school-based prevention curricula and mentoring programs. Prevention interventions have also been developed to enhance the skills of socialization agents such as parenting programs, programs that enhance instructional skills, or monitoring skills such as modeling and rewarding classroom or workplace behaviors.

Prevention Reflecting a Developmental Process

Prevention like other socialization agents reflects outcomes that are developmentally appropriate and address the competencies of the target individuals. Early or primary socialization of infants and children will have different expectations than that of later or secondary socialization. Fundamental to engaging in prosocial behaviors is the development of prosocial attitudes and beliefs, and these in turn grow out of emotional attachments to parents, family members, and members of an individual's community. Furthermore, in congruence with the developmental age and competencies of the targeted individuals, their participation in the socialization process whether through the socialization agents or prevention interventions may be more directive or supportive.

Prevention interventions can assist families and other socialization agents in achieving these goals or may specifically target them through formal program activities. The examples provided below demonstrate for a variety of prevention intervention approaches how the socialization process is applied to prevent targeted individuals from deciding on performing negative behaviors in different contexts.

Prevention as Evidence-Based Socialization: Applications

In the following, we apply our perspective of prevention as evidence-based socialization to three well-known and evidence-based programs, namely, the

Nurse–Family Partnership (NPF), the Good Behavior Game (GBG), and the Life Skills Training (LST).

Nurse–Family Partnership

NFP, originally developed by Olds (1988), is a selected family-based intervention serving low-income, at-risk pregnant women bearing their first child. NFP is designed to reduce an array of negative interdependent outcomes, which are concentrated in low-income families (i.e., poor birth outcomes, child abuse and neglect, welfare dependence, and poor maternal life course). In NFP, during pregnancy through the first two years of the child's life, nurses trained in women's and children's health serve as home visitors. During the home visits, these nurses promote healthy behavior related to pregnancy outcomes, support the mother in building positive relationships with family members and friends, and link women and their family members to additional health and human service agencies. NFP originally evaluated in a small, rural county in New York State was later replicated in Elmira (Kitzman et al., 1997) and in Denver and has shown to improve mother's health-related behaviors (e.g., substance use), to reduce dysfunctional caregiving (e.g., child abuse and neglect), and to improve maternal life course (e.g., welfare dependency) and antisocial behavior of the adolescent child (e.g., running away and delinquency).

NFP and the Integrated Prevention Science Model (IPSM). Through the integration of theories of human ecology and self-efficacy, NFP argues that the mothers' engagement in unhealthy behaviors is the result of observing and mimicking the inefficient behavior of prior role models (e.g., their own parents) in concert with a limited reservoir of successful experiences and a lack of information about positive health-related behaviors during pregnancy. These three important aspects of the decision-making process to engage in healthy behaviors during and after pregnancy (i.e., prior experiences and beliefs, personal beliefs about self-efficacy, and evidence-based information) form the targets of the nurse visitations. NFP thus is designed to socialize women to be better mothers by helping them understand the relationship between their own behavior and the health and development of their baby and to aid mothers in establishing realistic goals and achievable objectives, thus enriching their behavioral repertoire to be used in later situations. This training enhances the mothers' parenting skills and improves the socialization of their children reflected in their positive outcomes.

The Good Behavior Game

GBG, originally developed by Barrish, Saunders, and Wolf (1969), is a universal classroom-centered intervention supporting teachers in managing the classroom. In other words, GBG supports teachers as the educational field-specific agents in socializing elementary students into the role of the student, such as sitting still,

talking in turn, and paying attention. GBG is primarily a group-based behavior modification program, in which teams that display disruptive behaviors, such as verbal and physical disruptions and noncompliance, receive marks on the board. The team with the lowest number of marks is rewarded (Kellam et al., 2011). In addition to 20 small observational, randomized studies that showed short-term improvement in student behavior, GBG was evaluated in a large randomized trial of 2,311 first graders in the Baltimore Prevention trial (Kellam et al., 2008). The GBG intervention has been shown to have both short- and long-term impact on aggressive behavior, substance use, antisocial personality disorder, and criminality (Kellam et al., 2008; Kellam, Rebok, Ialongo, & Mayer, 1994; Petras et al., 2008) and has been replicated in a second trial in Baltimore (Ialongo, Poduska, Werthamer, & Kellam, 2001; Petras, Masyn, & Ialongo, 2011) as well as in the Netherlands (van Lier, Huizink, & Vuijk, 2011) and in Belgium (Leflot, van Lier, Onghena, & Colpin, 2010).

GBG and IPSM. From the IPSM perspective, GBG emphasizes that a child's negative life course occurs and is produced in the interaction with developmentally specific agents, such as the teacher. Specifically, the IPSM indicates that the engagement in aggressive/disruptive behavior, on the one hand, and a teacher's response to such behavior, on the other hand, are the end result of a decision-making process reflective of socialization that has occurred at an earlier point in time. It is well supported that children's level of aggressive/disruptive behaviors in elementary school is in part the result of parents' failure to effectively punishing noncompliant, aggressive, disruptive behavior during the toddler years, the first step in a process serving to train the child to become progressively more coercive and antisocial. When confronted with such children in the classroom, teachers have an array of responses at their disposal varying as a function of their training. Many teachers lack evidence-based skills for classroom management and tend to apply negative responses, such as school suspension, thus further escalating the teacher–student interaction. As such, GBG provides the teacher with tools to make decisions that are more effective in dealing with disruptive students, and, in turn, it teaches students behavioral alternatives for classroom behavior.

Life Skills Training

LST was developed by Botvin in the 1980s as a universal substance use prevention program originally designed to target adolescents when they were in the 7th grade. The initial program that was evaluated consisted of a core curriculum of 15 sessions followed with 10 booster sessions delivered when these students were in the 8th grade and five sessions when they were in the 9th grade (Botvin, Baker, Dusenbury, Botvin, & Diaz, 1995). Over the years, LST has been altered to include a 24-session elementary school program (that has not been evaluated), the 30-session middle school program, and a 10-session high school program. All of the programs focus on substance use and violence through increasing individual self-esteem, instilling prosocial attitudes, and providing life skills including personal self-management

skills, general social skills, and drug use resistance skills. The specifics of these program elements are designed to be age-appropriate. Generally, teachers, peer leaders and other trained adults deliver the program in schools.

LST and IPSM. LST and similar effective universal school-based prevention curricula build on the experiences and research of the 1970s on prevention programs designed to increase knowledge of substance use consequences (Evans et al., 1978), the early work on psychosocial development that intended to enhance personal and social development, and the emergence of Bandura's SLT (1977) and Jessor and Jessor's Problem Behavior Theory (1977). Botvin and his group conceptualized substance use as a socially learned behavior resulting from both social and personal influences. Within this framework, a preventive intervention would involve enhancing individual competencies and providing adolescents problem-specific knowledge and skills that are designed to strengthen prosocial attitudes and behaviors and their skills to resist pressures to initiate substance use. This approach has been found by Botvin and colleagues (2006) to transfer to violence and other risky behaviors. Through the establishment of prosocial norms and enhancing adolescents' competencies to assume age-appropriate roles, LST is an exemplar of prevention as a socialization agent per se (Griffin et al., 2006; Griffin, Botvin, Nichols, & Doyle, 2003).

Conclusions

In this chapter, we have proposed a comprehensive perspective of prevention science, developing a conceptual model that integrates primary socialization models with social information processing and decision-making. In this model, we have emphasized that the ultimate goal of socialization is to instill culturally appropriate norms and values resulting in prosocial, conforming behavior of individuals while ensuring the behavioral reproduction of the society's culture. Applying this perspective to prevention, we conclude that in the prevention process evidence-based socialization content is communicated. This communication can be indirect by training relevant socialization agents or direct by prevention agents subsuming the role of a socialization agent itself. The advantage of this perspective is that it redirects the focus of prevention theories as they address issues related to the socialization process. In addition, since socialization is a universal experience, which might only differ in its content and messaging modality, the applicability of this integrated theory is not confined by geographic and cultural boundaries.

A Devil's Advocate Perspective

Some have argued (e.g., Foxcroft, 2009) that focusing on the cognitive and experiential determinants of the decision-making process is not effective because the ultimate focus of prevention should be on the behavior. Commonly, this emphasis is

supported by referencing Gibson's work on environmental affordances (Gibson, 1977, 1979). Affordances describe environmental features that support the individual's actions and intentions. They are assumed to be independent of the individual's ability to recognize them but are always in relation to the actors and therefore dependent on their capabilities. For example, Washington, DC's subway system has been deliberately designed to reduce crime by omitting public restrooms, luggage lockers, and excess seating to avoid motivated offenders from lingering and assessing targets. In addition, train platforms feature high arched ceilings with few supporting columns to increase the chance for natural surveillance (Mair & Mair, 2003). This example as well as other applications of environmental modifications that have shown success is compatible with the conceptual model for prevention that is presented in this chapter. For affordances to stimulate or redirect behavior, they need to be recognized and interpreted by the individual (Norman, 1988; Zaff, 1995). Importantly, the extent to which these affordances are recognized and interpreted is likely influenced by past knowledge and experiences internalized during the individual's socialization process.

Not all available prevention interventions incorporate a decision-making process or skills development (e.g., keg registration or carding purchasers of tobacco and alcohol). Rather than being skilled based, keg registration or carding purchasers of tobacco and alcohol may reinforce the norm against underage use of these substances but also take the decision-making regarding the use of alcohol or tobacco out of the hands of individuals by presenting a barrier or extra hurdle to deter use. These policy or environmental preventive interventions play a major role in preventing the inappropriate use of these substances, are generally population-based, and are presumably inexpensive to carry out. However, they require the training and active monitoring of the dispensers of these products and ultimately do not provide the required information or skills needed by those targeted by the intervention to internalize the concept that the use of alcohol and tobacco among children and adolescents is unhealthy and often associated with many negative outcomes. From an integrated perspective, it can then be argued that not considering relevant socialization agents responsible for promoting norms and skills for healthy behavior and their impact on individual decision-making will likely diminish the effectiveness of purely environmental interventions.

On the other hand, the criticism against individual/group behavioral interventions such as school-based prevention curricula or even parenting skills training is that unless the programs are sustained over long periods, thus involving multiple generations, they will not have a population impact. Logically then, there is a need for both types of interventions to reinforce the desired behavioral norms of the community and increase the likelihood of success (Flay, 2009). Ultimately, the issue for communities to address is not environmental versus individual interventions but how best to match the targeted outcomes and messages of both (Hawkins et al., 2012).

References

Ajzen, I. (1988). *Attitudes, personality, and behavior*. Chicago, IL: Dorsey Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1977). *Social learning theory*. New York, NY: General Learning Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

Barrish, H., Saunders, M., & Wolf, M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*(2), 119–124.

Biglan, A., Domitrovich, C., Ernst, J., Etz, K., Mason, M. J., Robertson, E., et al. (2011). *Standards of knowledge*. Fairfax, VA: Society for Prevention Research.

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*(14), 1106–1112.

Botvin, G. J., Griffin, K. W., & Nichols, T. R. (2006). Preventing youth violence and delinquency through a universal school-based prevention approach. *Prevention Science, 7*, 403–408.

Brook, J. S., Brook, D. W., Gordon, A. S., Whiteman, M., & Cohen, P. (1990). The psychosocial etiology of adolescent drug use: a family interactional approach. *Genetic, Social, and General Psychology Monograph, 116*(2), 111–267.

Casey, B. J., Jones, R. M., Levita, L., Libby, V., Pattwell, S., Ruberry, E., et al. (2010). The storm and stress of adolescence: Insights from human imaging and mouse genetics. *Developmental Psychobiology, 52*(3), 225–235.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Annals of the American Academy of Political and Social Science, 591*, 98–124.

Catalano, R. F., Fagan, A. A., Gavin, L. E., Greenberg, M. T., Irwin, C. E., Ross, D. A., et al. (2012). Worldwide application of prevention science in adolescent health. *Lancet, 379*, 1653–1664.

Catalano, R. F., Hawkins, J. D., Berglund, M. L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*(6 Supplement), 230–239.

Catalano, R. F., Kosterman, R., Hawkins, J. D., Newcomb, M. D., & Abbott, R. D. (1996). Modeling the etiology of adolescent substance use: A test of the social development model. *Journal of Drug Issues, 26*, 429–455.

Chang, L. J., & Sanfey, A. G. (2008). Emotion, decision-making and the brain. *Advances in Health Economics and Health Services Research, 20*, 31–53.

Cicchetti, D., & Blender, J. A. (2006). A multi-level-of-analysis perspective on resilience: Implications for the developing brain, neural plasticity, and preventive interventions. *Annals of the NY Academy of Science, 1094*, 248–258.

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., et al. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist, 48*(10), 1013–1022.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74–101.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York, NY: Avon Books.

Dodge, K. A., & Crick, N. R. (1990). Social information-processing bases of aggressive behavior in children. *Personality and Social Psychology Bulletin, 16*(1), 8–22.

Evans, R. I., Rozelle, R. M., Mittlemark, M. B., Hansen, W. B., Bane, A. L., & Havis, J. (1978). Deterring the onset of smoking in children: Knowledge of immediate physiological effects and coping with peer pressure, media pressure, and parent modeling. *Journal of Applied Social Psychology, 8*, 126–135.

Fareri, D. S., Martin, L. N., & Delgado, M. R. (2008). Reward-related processing in the human brain: Developmental considerations. *Developmental Psychopathology, 20*, 1191–1211.

Fishbein, M. (2000). The role of theory in HIV prevention. *AIDS Care, 12*, 273–278.

Fishbein, M. (2008). A reasoned action approach to health promotion. *Medical Decision Making, 28*, 834–844.

Flay, B. R. (2009). School-based smoking prevention programs with the promise of long-term effects. *Tobacco Induced Diseases, 5*(1), 7.

Foxcroft, D. R. (2009). *Classifying prevention: Form, function and theory*. Presentation at the 2nd Annual European Society for Prevention Research, Lisbon, Portugal.

Garmezy, N. (1985). Broadening research on developmental risk. In W. Frankenburg, R. Emde, & J. Sullivan (Eds.), *Early identification of children at risk: An international perspective* (pp. 289–303). New York, NY: Plenum.

Garmezy, N. (1991). Resiliency and vulnerability to adverse developmental outcomes associated with poverty. *American Behavioral Scientist, 34*(4), 417–429.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing* (pp. 67–82). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton-Mifflin.

Glantz, M. D., & Pickens, R. W. (1992). Vulnerability to drug abuse: Introduction and overview. In M. D. Glantz & R. W. Pickens (Eds.), *Vulnerability to drug abuse* (pp. 1–14). Washington, DC: American Psychological Association Books.

Graber, J. A., Hill, J., & Saczawa, M. (2013). Childhood and the entry into adolescence: A pivotal period in health-related behaviors and prevention. In Z. Sloboda & H. Petras (Eds.), *Advances in prevention science: Defining prevention science*. New York, NY: Springer Publication.

Griffin, K. W., Botvin, G. J., & Nichols, T. R. (2006). Effects of a school-based drug abuse prevention program for adolescents on HIV risk behavior in young adulthood. *Prevention Science, 7*(1), 103–112.

Griffin, K. W., Botvin, G. J., Nichols, T. R., & Doyle, M. M. (2003). Effectiveness of a universal drug abuse prevention approach for youth at high risk for substance use initiation. *Preventive Medicine, 36*, 1–7.

Grusec, J. E., & Davidov, M. (2010). Integrating different perspectives on socialization theory and research: A domain-specific approach. *Child Development, 81*, 687–709.

Guerra, N. G., Nucci, L., & Huesmann, L. R. (1994). Moral cognition and childhood aggression. In L. R. Huesmann (Ed.), *Aggressive behavior: Current perspective* (pp. 13–33). New York, NY: Plenum Press.

Haddon, W. (1972). A logical framework for categorizing highway safety phenomena and activity. *Journal of Trauma, 12*, 193–207.

Hawkins, J. D., Catalano, R. F., & Miller, J. Y. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: Implications for substance abuse prevention. *Psychological Bulletin, 112*, 64–105.

Hawkins, J. D., Oesterle, S., Brown, E. C., Monahan, K. C., Abbott, R. D., Arthur, M. W., et al. (2012). Sustained decreases in risk exposure and youth problem behaviors after installation of the communities that care prevention system in a randomized trial. *Archives of Pediatric and Adolescent Medicine, 166*, 141–148.

Herrenkohl, T. I., Maguin, E., Hill, K. G., Hawkins, J. D., Abbott, R. D., & Catalano, R. F. (2000). Developmental risk factors for youth violence. *Journal of Adolescent Health, 26*(3), 176–186.

Ialongo, N., Poduska, J., Werthamer, L., & Kellam, S. (2001). The distal impact of two first-grade preventive interventions on the conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioral Disorders, 9*, 146–160.

Israelashvili, M. (2012). Should adjustment disorder be conceptualized as transitional disorder? In pursuit of adjustment disorders definition. *Journal of Mental Health, 6*, 579–588.

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education Quarterly, 1*, 1–47.

Jensen, P. S., Martin, D., & Cantwell, D. P. (1997). Comorbidity in ADHD: Implications for research, practice, and DSM-V. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1065–1079.

Jessor, R., & Jessor, S. L. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth*. New York, NY: Academic.

Jonah, B. A. (1997). Sensation seeking and risky driving: A review and synthesis of the literature. *Accident Analysis and Prevention, 5*, 651–665.

Kellam, S. G., Branch, J. D., Agrawal, K. C., & Ensminger, M. E. (1975). *Mental health and going to school: The Woodlawn Program of assessment, early intervention, and evaluation*. Chicago, IL: University of Chicago Press.

Kellam, S. G., Brown, C. H., Poduska, J. M., Ialongo, N. S., Wang, W., Toyinbo, P., et al. (2008). Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence, 95* (Suppl 1), S5–S28.

Kellam, S. G., Mackenzie, A. C., Brown, C. H., Poduska, J. M., Wang, W., Petras, H., et al. (2011). The good behavior game and the future of prevention and treatment. *Addiction Science and Clinical Practice, 6*(1), 73–84.

Kellam, S. G., Rebok, G. W., Ialongo, N. S., & Mayer, L. S. (1994). The course and malleability of aggressive behavior from early first grade into middle school: Results of a developmental epidemiologically-based preventive trial. *Journal of Child Psychology and Psychiatry, 35*, 259–282.

Kim-Spoon, J., & Farley, J. P. (2013). Individual physical, cognitive, and emotional vulnerability across the life course: Benchmarks and developmental challenges in adolescence and early adulthood. In Z. Sloboda & H. Petras (Eds.), *Advances in prevention science: Defining prevention science*. New York, NY: Springer Press.

Kitzman, H., Olds, D. L., Henderson, C. R., Jr., Hanks, C., Cole, R., Tatelbaum, R., et al. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. *Journal of the American Medical Association, 278*(8), 644–652.

Kuczynski, L. (2003). Beyond bidirectionality: Bilateral conceptual frameworks for understanding dynamics in parent-child relations. In L. Kuczynski (Ed.), *Handbook of dynamics in parent-child relations* (pp. 1–24). Thousand Oaks, CA: Sage.

Lacourse, E., Cote, S., Nagin, D., Vitaro, F., Brendgen, M., & Tremblay, R. (2002). A longitudinal-experimental approach to testing theories of antisocial behavior development. *Development and Psychopathology, 14*, 909–924.

Lahey, B. B., McBurnett, K., & Loeber, R. (2000). Are attention-deficit/hyperactivity disorder and oppositional defiant disorder developmental precursors to conduct disorder? In A. J. Sameroff & M. Lewis (Eds.), *Handbook of developmental psychopathology* (2nd ed., pp. 431–446). New York, NY: Plenum Press.

Leflot, G., van Lier, P. A., Onghena, P., & Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Abnormal Child Psychology, 6*, 869–882.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development, 71*, 107–118.

Leukefeld, C. G., & Leukefeld, S. (1999). Primary socialization theory and a bio/psycho/social/spiritual practice model for substance use. *Substance Use & Misuse, 34*, 983–991.

Loeber, R., Wung, P., Keenan, K., Giroux, B., Stouthamer-Loeber, M., Van Kammen, W. B., et al. (1993). Developmental pathways in disruptive child behavior. *Development and Psychopathology, 5*, 103–133.

Mair, J. S., & Mair, M. (2003). Violence prevention and control through environmental modifications. *Annual Review of Public Health, 24*, 209–225.

Masten, A., Best, K., & Garmezy, N. (1990). Resilience and development: Contributions from the study of children who overcame adversity. *Development and Psychopathology, 2*, 425–444.

Masten, A. S., Burt, K., & Coatsworth, J. D. (2006). Competence and psycho-pathology in development. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology risk, disorder, and adaptation* (Vol. 3, pp. 696–738). New York, NY: Wiley.

Moffitt, T. E. (1993). Adolescent-limited and life-course-persistent, Antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674–701.

Mrazek, P. J., & Haggerty, R. J. (Eds.). (1994). *Reducing risks for mental disorders: frontiers for preventive mental disorders*. Washington, DC: National Academy Press.

Norman, D. A. (1988). *The design of everyday things*. New York, NY: Doubleday.

O'Connell, M. E., Boat, T., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academy Press.

Oetting, E. R. (1999). Primary socialization theory. Developmental stages, spirituality, government institutions, sensation seeking and theoretical implications. *Substance Use and Misuse, 7*, 947–982.

Oetting, E. R., Deffenbacher, J. L., & Donnermeyer, J. F. (1998). Primary socialization theory: The role played by personal traits in the etiology of drug use and deviance, II. *Substance Use and Misuse, 33*(6), 1337–1366.

Oetting, E. R., & Donnermeyer, J. F. (1998). Primary socialization theory: The etiology of drug use and deviance, I. *Substance Use and Misuse, 33*(4), 975–1026.

Oetting, E. R., Donnermeyer, J. F., & Deffenbacher, J. L. (1998). Primary socialization theory. The influence of the community on drug use and deviance, III. *Substance Use & Misuse, 33*(8), 1629–1665.

Oetting, E. R., Donnermeyer, J. F., Trimble, J. E., & Beauvais, F. (1998). Primary socialization theory: Culture, ethnicity, and cultural identification—The links between culture and substance use, IV. *Substance Use and Misuse, 33*(10), 2075–2107.

Olds, D. (1988). The prenatal/early infancy project. In R. Price, E. Cowen, R. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A casebook for practitioners*. Washington, DC: American Psychological Association.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*, 329–335.

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *A social learning approach, IV: Antisocial boys*. Eugene, OR: Catalia.

Petraitis, J., Flay, B., & Miller, T. Q. (1995). Reviewing theories of adolescent substance use: Organizing pieces in the puzzle. *Psychological Bulletin, 1*, 67–86.

Petras, H., Kellam, S. G., Brown, C. H., Muthén, B. O., Ialongo, N. S., Poduska, J. M., et al. (2008). Developmental epidemiological courses leading to antisocial personality disorder and violent and criminal behavior: Effects by young adulthood of a universal preventive intervention in first- and second-grade classrooms. *Drug and Alcohol Dependence, 95*(Suppl 1), S45–S59.

Petras, H., Masyn, K., & Ialongo, N. (2011). The developmental impact of two first grade preventive interventions on aggressive/disruptive behavior in childhood and adolescence: An application of Latent Transition Growth Mixture Modeling. *Prevention Science, 12*, 300–313.

Poldrack, R. A., & Packard, M. G. (2003). Competition between memory systems: Converging evidence from animal and human studies. *Neuropsychologia, 41*, 245–251.

Rauh, H., & Bergmann, K. E. (2013). Physical, cognitive, and emotional vulnerability across the life course: Benchmarks and developmental challenges during the prenatal and infancy period. In Z. Sloboda & H. Petras (Eds.), *Advances in prevention science: Defining prevention science*. New York, NY: Springer Press.

Rebok, G. W., Parisi, J. M., & Kueider, A. M. (2013). Stressors and vulnerabilities in middle and old age: Opportunities for prevention. In Z. Sloboda & H. Petras (Eds.), *Advances in prevention science: Defining prevention science*. New York, NY: Springer Press.

Redmond, C., Spoth, R. L., Shin, C., Schainker, L. M., Greenberg, M. T., & Feinberg, M. (2009). Long-term protective factor outcomes of evidence-based interventions implemented by community teams through a community-university partnership. *Journal of Primary Prevention, 30*, 513–530.

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. Cacioppo & R. Petty (Eds.), *Social psychophysiology*. New York, NY: Guilford Press.

Romer, D., & Hennessy, M. (2007). A biosocial-affect model of adolescent sensation seeking: The role of affect evaluation and peer-group influence in adolescent drug use. *Prevention Science, 8*, 89–101.

Romer, D., Sznitman, S. R., Hennessy, M., DiClemente, R. J., Salazar, L. F., Vanable, P. A., et al. (2009). Mass media as an HIV-prevention strategy: Using culturally sensitive messages to reduce HIV-associated sexual behavior of high risk African-American youth. *American Journal of Public Health, 12*, 2150–2159.

Rutter, M. (1985). Resilience in the face of adversity. Protective factors and resistance to psychiatric disorder. *British Journal of Psychiatry, 147*, 598–611.

Rutter, M., Dunn, J., Plomin, R., Simonoff, E., Pickles, A., Maughan, B., et al. (1997). Integrating nature and nurture: Implications of person-environment correlations and interactions for developmental psychopathology. *Development and Psychopathology, 9*(2), 335–364.

Satterfield, J. H., & Schell, A. (1997). A prospective study of hyperactive boys with conduct problems and normal boys: Adolescent and adult criminality. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1726–1735.

Sloboda, Z., Glantz, M. D., & Tarter, R. E. (2012). Revisiting the concept of risk and protective factors for understanding the etiology and development of drug abuse: Implications for prevention. *Substance Use and Misuse, 47*, 944–962.

Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Developmental Psychobiology, 52*, 216–224.

van Lier, P. A., Huizink, A., & Vuijk, P. (2011). The role of friends' disruptive behavior in the development of children's tobacco experimentation: Results from a preventive intervention study. *Journal of Abnormal Child Psychology, 39*(1), 45–57.

Werner, E. (1989). High-risk children in young adulthood: A longitudinal study from birth to 32 years. *American Journal of Orthopsychiatry, 59*(1), 72–81.

Zaff, B. S. (1995). Designing with affordances in mind. In J. M. Flach, P. A. Hancock, J. Caird, & K. J. Vicente (Eds.), *Global perspectives on the ecology of human–machine systems* (Resources for ecological psychology, Vol. 1, pp. 238–272). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zuckerman, M. (1994). *Behavioural expressions and biosocial bases of sensation-seeking*. Cambridge: Cambridge University Press.

Chapter 12
Design of Prevention Interventions

Louise A. Rohrbach

Introduction

The design of disease prevention and health promotion interventions is a complex process that involves careful decision-making about the health behaviors one aims to change or other outcomes one aims to achieve. This includes understanding the etiologic mechanisms associated with the behavior, such as the prevalence and the risk and protective factors that contribute to it; the intervention strategies that will be effective in producing behavior change; the population subgroups to target with these strategies; the most appropriate modes of program delivery and implementation; how to take into account the developmental stage and other characteristics of the target population; and other issues. A key task in intervention design is to develop and apply a thorough understanding of theory on the determinants of the target health behaviors or other outcomes as well as how to affect change in those behaviors and other outcomes. For the prevention science field to achieve its ultimate goal of improving public health, there must be ongoing development and testing of interventions that build on the latest basic scientific findings, methods, and theory. This chapter provides an overview of the intervention design process and describes three key stages in the process, including the application of a theoretical framework, building the intervention, and pilot testing the intervention.

Interventions may be defined as programs, policies, practices, and guidelines encompassing intentional actions (singular or constellation of actions) designed for an individual, organization, community, region, or system that are intended to alter health behaviors, address risk or protective factors, and improve health-related outcomes (Centers for Disease Control and Prevention, 2007a; Rabin, Brownson, Haire-Joshu, Kreuter, & Weaver, 2008). Applying this definition, a prevention intervention might be as explicit as a program with a teacher's manual that outlines

L.A. Rohrbach (✉)
Preventive Medicine, University of Southern California, 2001 N. Soto St., Room 302 K, Los Angeles, CA 90032, USA
e-mail: rohrbac@usc.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_12,

all instructional steps, content points, and discussion questions to be delivered (e.g., Life Skills Training, Botvin, Baker, Dusenbury, Tortu, & Botvin, 1990; Botvin, Mihalic, & Grotpeter, 1998) or as broad as a set of guidelines for prevention efforts (e.g., the CDC manual on best practices for state-level tobacco control programs, CDC, 2007b). Furthermore, an intervention may consist of a single method or strategy such as the Life Skills Training program—a substance abuse prevention curriculum designed for implementation in schools—or it may comprise a combination of strategies designed to work together synergistically, such as state tobacco control programs that deliver antitobacco media, implement school-based tobacco prevention curricula, and initiate smoke-free policies at the community and state levels (e.g., Rohrbach et al., 2002). The primary target for intervention activities might be individuals in specific settings (e.g., school, clinics, communities), organizations (e.g., policies and practices of the organization and the individuals that put them into practice), or systems (e.g., states, regions, and social networks). Finally, effective disease prevention and health promotion interventions are based on a theory of change that specifies the process by which changes in outcomes will occur (Bartholomew & Mullen, 2011; Issel, 2009).

To provide a context for intervention design, it is useful to consider where it occurs within the life cycle of biomedical research, from basic to applied science (Greenwald, 1990). The life cycle of a prevention intervention is conceptualized as a five-phase process: (1) research on the epidemiology of the target problem or disorder; (2) research on the etiology of the problem, including the risk and protective factors for the problem that may be potential targets for the preventive intervention; (3) design, pilot testing, and efficacy trials of the intervention (i.e., studies of whether the intervention works when implemented under optimum conditions); (4) effectiveness testing of the intervention (i.e., studies of whether the intervention works when delivered under real-world conditions); and (5) dissemination and implementation trials (Mrazek & Haggerty, 1994). Dissemination research encompasses a broad range of topics, such as the generalizability of the program and translation research that investigates the complex processes and systems through which evidence-based interventions are adopted, implemented, and sustained on a large scale (Spoth et al., 2013). Thus, intervention design involves the application of etiological research findings from the earliest phases in the research cycle and leads to studies that will allow the researcher to evaluate the efficacy, effectiveness, and dissemination potential of the intervention.

Stages of the Intervention Design Process

Several models of the process of intervention design have been posited, such as Worden's social marketing/formative evaluation approach (Worden et al., 1988) and Sussman's models of empirical curriculum and program development (Sussman, 1991; Sussman & Wills, 2001). In the public health field, the most widely used model is Precede-Proceed (Green & Kreuter, 2005), which describes

the entire cycle of program planning from needs assessment and development to implementation and evaluation, with feedback loops that connect each phase of the cycle. These models propose that intervention development occurs in a series of linked steps or stages. Synthesizing across the models, the key stages that occur before program implementation include (1) adopting a theoretical foundation to guide the research, (2) building the program, and (3) pilot testing. After the intervention is designed and ready for implementation, the life cycle continues with research to test its efficacy, effectiveness, and dissemination. In the next section of the chapter, the intervention development steps are discussed in detail.

Before beginning the intervention design process, one needs to develop a design team comprising individuals and organizations with complementary, but distinct, sets of skills and expertise. In efficacy research, often the team is small, comprising researchers and representatives of stakeholder groups that will ultimately participate in and use the intervention. In efficacy and dissemination research, usually there is a broader range of stakeholders involved in the design team, including evaluators, administrators of organizations, program providers, and consumers.

Much has been written about the value of adopting a transdisciplinary approach to understanding health and social problems and potential solutions to them (Rosenfeld, 2002). Prevention science is inherently transdisciplinary, with contributions from researchers in psychology, sociology, public health, medicine, education, and other disciplines. However, while the scientific approach to designing prevention interventions generally has been transdisciplinary, many have argued that the design process has not adequately incorporated the practitioner perspective, which has contributed to a gap between interventions found to be efficacious and those that are ready for wide-scale dissemination and implementation (Rohrbach, D'Onofrio, Backer, & Montgomery, 1996; Rohrbach, Grana, Sussman, & Valente, 2006). Thus, to help ensure that an intervention will progress through the life cycle from epidemiological research to dissemination and address the priorities of its participants and users, it is important to obtain and apply input from a variety of perspectives during the design process, such as researchers, practitioners, and members of the intervention target group (Hodges & Videto, 2011).

Stage 1: Adopting a Theoretical Foundation

Identifying the Target Group

The intervention design process should begin with a thorough review of empirical research on the nature and extent of the health problem or disorder to be prevented, including rates of incidence and prevalence of the problem by population subgroups as well as data on the age of onset of the problem. One purpose of reviewing epidemiological data is to guide the selection of the target group for the intervention, that is, the demographic group most in need of the program or most critical for intervention based on their substantial prevalence of the problem and/or risk for

increasing the prevalence. High-risk groups targeted for prevention programs can be identified at multiple levels of analysis, including individuals, families, and communities (National Research Council and Institute of Medicine, 2009).

In the health and mental health fields, interventions have been classified on the basis of the population that is targeted. In public health contexts, prevention interventions are described as primary, secondary, or tertiary. Primary prevention is designed to prevent a disease from beginning, secondary prevention involves efforts to prevent manifestation of disease among those at elevated risk, and tertiary prevention involves activities to limit the extent of existing disease (Issel, 2009). In 1983, Gordon proposed an alternative classification that is more relevant to behavioral health (Gordon, 1983), which was adopted by the IOM and has become the predominant framework for conceptualizing prevention of mental, emotional, and behavioral disorders (Mrazek & Haggerty, 1994). The IOM categories of prevention interventions include (1) universal prevention, which are strategies that can be implemented with the entire population; (2) selected prevention, which refers to strategies that are targeted to a subpopulation identified as being at elevated risk for a disorder; and (3) indicated prevention, which includes strategies that are targeted to individuals identified as having an increased vulnerability to a disorder based on some individual assessment.

Decisions about the type of prevention approach to adopt take into account the costs and benefits of delivering the intervention to the targeted population and provide a framework for evaluating intervention effectiveness (National Research Council and Institute of Medicine, 2009). For example, universal prevention approaches are likely to provide some benefit to all participants, with relatively little cost or risk of negative consequences for them. On the other hand, selective and indicated prevention strategies often involve more intensive activities that may produce greater costs to participants in terms of time and effort, yet the potential benefit from participation may be greater. The success of prevention interventions will be evaluated in terms of reductions in the probability of developing a specific disorder and the extent to which these reductions are sustained over time, but the relative emphasis placed on specific outcome indicators (e.g., behaviors that increase risk of developing the disorder, consequences of risk behaviors, evidence of early symptoms) may vary by the type of prevention approach that is adopted (National Research Council and Institute of Medicine, 2009).

The Use of Theory

In intervention research, the application of theory and empirical evidence not only provides an understanding of the environmental and behavioral determinants of behavior related to a specified health problem, but it also explicates potential mechanisms for producing change in the outcome or behavior of interest and informs the selection of intervention methods to achieve those changes (Bartholomew & Mullen, 2011). The concept of risk and protective factors is often applied to develop an understanding of determinants of problem outcomes.

Over the past several decades, a substantial literature has emerged on the risk and protective factors for numerous mental, emotional, and behavioral disorders among young people (National Research Council and Institute of Medicine, 2009; also see Chap. 11 in this book). Risk factors, or measurable characteristics of an individual that precede and are associated with an outcome, can occur at multiple levels of the social ecology, including biological, psychological, family, community, and cultural levels. Protective factors are those characteristics at the individual, family, or community level that are associated with a lower likelihood of a problem outcome. In the early stages of intervention design, researchers often conduct what is sometimes referred to as "needs assessment" (Hodges & Videto, 2011), which may include both a review of the literature on risk and protective factors for the problem outcome and the collection of primary data on the prevalence of these factors among the selected target group.

Based on the findings of the needs assessment, the design team adopts a theoretical or "logic" model that identifies which determinants (risk and protective factors) are amenable to change by the intervention activities and specifies hypotheses about how these factors will interact over time to lead to the outcome. Theories from social psychology, health behavior, communications, sociology, social ecology, and other disciplines serve as a framework for explaining why health-compromising behaviors occur and how they might be improved (cf Glanz, Rimer, & Viswanath, 2008). For example, many studies of the relationships between risk and protective factors and problem behaviors among youth have been guided by, and provided evidence for, social psychological theories such as social learning theory (Bandura, 1977), problem behavior theory (Jessor & Jessor, 1977), and the theory of planned behavior (Ajzen, 1991). In addition, a number of ecological models have been posited, which delineate factors from multiple levels of the human ecology (e.g., intrapersonal, sociocultural, policy, and physical–environmental) that are likely to interact to influence health-related behaviors (Sallis, Owen, & Fisher, 2008).

Overall, a key issue for the prevention science field is to determine which risk and protective factors are actually changed by a specific intervention and whether changes in those factors mediate, or account for, changes in the problem outcome. Randomized control trials of prevention interventions are viewed as the most rigorous methods to provide evidence in support of or counter to specific theoretical concepts as mediators of change in risk behaviors (Coie et al., 1993). Even when quasi-experimental designs are used, intervention researchers can help advance the science of behavior change by investigating and reporting what constructs mediate intervention effects. Bartholomew and Mullen (2011) have suggested that effective development and testing of prevention interventions often involve an approach to theory that is fundamentally different from generating theory, testing single theories, or comparing the effectiveness of one or more theories for explaining or changing behavior. In the design of interventions, prevention scientists tend to apply multiple theories and empirical evidence, and the theories for determinants of the problem may differ from the theories for its solution. The main focus of intervention is on promoting behavior change and improving health outcomes, and

the criteria for success usually are formulated in terms of the problem one is trying to solve rather than the theory that may serve as the foundation for the intervention. Thus, the contributions to theory made by intervention research tend to focus on the problem-solving process.

Applying Theory in Interventions: The Example of Social Norm Change

School-based substance abuse prevention programs that have been adopted widely, such as Life Skills Training (Botvin et al., 1990, 1998) and Project ALERT (Ellickson & Bell, 1990), provide examples of how social psychological theories have been applied in intervention design. A key component of these programs focuses on changing youths' social normative beliefs about substance use. According to the theory of reasoned action (Ajzen & Fishbein, 1980), one's behavior is influenced by normative beliefs or the perceived approval or disapproval of the behavior in question by key people in one's social environment. Empirical research has indicated that young people tend to overestimate the prevalence of all forms of substance use, and expectations of substance use among peers are positively associated with self-reported substance use (Botvin, Botvin, Baker, Dusenbury, & Goldberg, 1992; Sussman et al., 1988; Wolfson, 2000). Given this empirical and theoretical foundation, prevention program lessons were designed with the aim of correcting young people's erroneous perceptions about the prevalence and acceptance of substance use among peers. Examples of specific activities include presentation and discussion of national data on substance use prevalence, administering and summarizing local surveys on substance use, and using Socratic questioning and debate methods to clarify group expectations regarding what is appropriate and desirable with regard to substance use (Ellickson & Bell, 1990; Hansen & Graham, 1991; Rohrbach, English, Hansen, & Johnson, 2000).

Redefinition of social norms of sexual behavior toward avoidance of high-risk practices has also been a focus of effective HIV prevention programs targeting homosexual men (Catania et al., 1991; Kegeles, Hays, & Coates, 1996; Kelly et al., 1997). In one intervention that demonstrated significant reductions in several risk indicators, popular, well-liked members of social networks (i.e., opinion leaders) were trained to engage in conversations with their peers that endorsed and supported risk avoidance (Kelly et al., 1997). In a similar approach, trained peers conducted outreach to friends and acquaintances both to promote safer sex messages and to encourage participation in intervention events (Kegeles et al., 1996).

Developing Intervention Objectives

An important element of the theoretical perspective for the prevention intervention is developing behavioral objectives that specify the health behaviors that will be the focus of efforts to produce change. For example, some interventions focused on reducing unwanted pregnancy and sexually transmitted infections among

adolescents have aimed to delay the onset of sexual intercourse by encouraging abstinence. An alternative model, which is considered a "harm reduction" approach, has both encouraged abstinence and provided instruction about "safe sex" behaviors such as the use of condoms and other contraception (Scher, Maynard, & Stagner, 2006). Thus, an intervention may be designed to delay or prevent a specific behavior (e.g., sexual intercourse), promote a specific behavior (e.g., condom use), or be a synthesis of both approaches. Generally, objectives related to health promotion are not considered to be separate from, but rather, one component of the spectrum of intervention objectives for both prevention and treatment of disorders (National Research Council and Institute of Medicine, 2009).

In sum, by the end of the first phase, the intervention design team will have identified the target group for the intervention, the type of intervention approach that will be used (i.e., universal, selective, or indicated), and theoretical foundation for the intervention. The design team may find it useful to develop a "logic model" or graphic representation of the hypothesized causal relationships between the determinants of the health problem that will be the focus of the intervention (i.e., mechanisms of change), the intervention activities that will address those determinants, and the expected behavioral and health outcomes (Bartholomew & Mullen, 2011; W. K. Kellogg Foundation, 2004). In the next phase of intervention design, the team begins to specify what the intervention activities will involve.

Stage 2: Building the Intervention

Methods and Strategies

In the prevention literature, a range of terms have been used to describe the building blocks or elements of interventions. Bartholomew, Parcel, Kok, and Gottlieb (2006) distinguish between *methods*, which are theory-informed processes for influencing change in the determinants of behavior and environmental conditions, and *strategies*, which are practical applications of the theoretical approach. For example, social learning theory-derived methods for increasing self-efficacy to change behavior might include modeling, skill training, guided practice with feedback, and reinforcement. Strategies for applying these methods might include step-by-step modeling and demonstration of skills in videotaped segments and in-person role-playing of common situations in which one would use the skill, followed by feedback and reinforcement. Other terms that have been used to describe the smallest identifiable content or instructional method of interventions are "component," "instructional method," and "activity" (Perry, 1999; Sussman & Wills, 2001).

The choice of methods for the intervention will be guided by the theoretical approach that has been adopted (Bartholomew et al., 2006). The intervention may use methods from one or several levels of the social ecology (intrapersonal, interpersonal, organizational/institutional, community, and public policy) to

address the risk and protective factors and environmental conditions that will be the focus of change by the intervention. For example, a community-level method of policy change may be employed to reduce the availability of alcohol in communities, a risk factor for alcohol misuse among youth. An individual-level approach such as school-based alcohol prevention education may be developed to address the risk factors of alcohol-related attitudes, perceptions of peer norms regarding the prevalence and acceptability of alcohol use, and self-efficacy to resist offers of alcohol from peers. Some methods at the higher system levels may have a direct or mediated effect on lower (embedded) levels. For example, a media campaign may have a direct effect on adolescents by producing reductions in their use of alcohol, or it may reduce adolescent alcohol use indirectly by influencing a change in public policies regarding alcohol use in public places. Overall, the task of the design team is to investigate the range of theoretical methods that focus on individual behavior change as well as those designed to change systems (i.e., family, organization, community) and select methods that are likely to be most effective in changing the target risk and protective factors.

Typically, the design team aims to develop an intervention that not only builds on activities and strategies that have been shown to be effective in producing behavior change but also introduces novel activities that will be effective. A broad range of strategies, instructional methods, and activities are available for consideration in the design of prevention interventions, such as advocacy, community mobilization, group discussions, debates, educational games, lectures, media, self-appraisals, cooperative learning, and others (Hodges & Videto, 2011). The process of selecting strategies often begins with a systematic review of other interventions related to the program one is developing (D'Onofrio, 2001). Examples of sources of information on interventions include registries of evidence-based programs (e.g., Drug Free Schools Program & United States Department of Education, 2001; Elliott, 1997; National Institute on Drug Abuse, 2003; Substance Abuse and Mental Health Services Administration, 2007) and meta-analyses of intervention effects (e.g., Cochrane Collaboration, 2012). At a minimum, a systematic review provides an understanding of what types of intervention components have produced positive changes in the problem outcome. If an intervention approach has been tested with a variety of target populations, the available evidence will provide guidance about its generalizability and potential effectiveness for the target group that has been selected. A review of the intervention literature can lead to a "pool of activities" that the program design team can consider adapting for use in its own intervention (Sussman & Wills, 2001). In addition, the review can help stimulate ideas for new activities that the team hypothesizes will be effective to counter specific determinants of behavior.

Formative Research

To narrow down the pool of potential intervention activities under consideration for inclusion in the intervention, several types of formative studies can be conducted.

One of the most commonly used formative research methods is the focus group, in which a convenience sample of target group members is asked open-ended questions to determine such factors as interest, comprehensibility, perceived effectiveness, and acceptance of the potential program activities (Sussman & Wills, 2001). Focus group data might help generate ideas for real-world case studies and scenarios used in a health promotion curriculum, for example. Intervention developers also use focus groups extensively to obtain target audience feedback about new materials (e.g., print media, DVDs, public service announcements) that have been developed specifically for the intervention.

Another formative evaluation approach is to assess the acceptability of potential program activities by administering paper-and-pencil surveys and interviews to target group members. For example, a sample of potential participants might complete a survey that contains descriptions of the intervention activities under consideration, followed by questions that measure the perceived effectiveness and interest of each. To complement the information garnered from target group members, the design team can obtain input from practitioners who are knowledgeable about the target group, such as classroom teachers or medical practitioners, by employing Delphi techniques via surveys, interviews, or face-to-face discussion groups (Sussman & Wills, 2001).

Target Group Characteristics

In making choices about the types of intervention activities one will implement, it is important to consider characteristics of the target group, such as age, gender, ethnicity, socioeconomic status, and cultural background. For example, theories in the fields of education and psychology posit that students differ in learning styles, based on cognitive, affective, and physiological factors (Keefe, 1987). Thus, while cooperative learning activities may be most effective with some students, self-appraisals or journaling may work best for others. Further, it has been hypothesized that the life experiences of adults affect their motivation and approach to learning (e.g., Knowles, 1978), suggesting the need for careful selection of intervention strategies when parents or other adults are the target group for prevention.

In addition to developmental appropriateness, the design team needs to consider which activities will be appropriate for the cultures represented in the target groups and reflect the core values, beliefs, norms, and other aspects of their worldviews and lifestyles (Gilbert, Sawyer, & McNeill, 2011; Resnikow, Baranowski, Ahluwalia, & Braithwaite, 1999). As an example, prevention interventions that use the arts to convey key messages may be very appropriate for use with groups whose cultural traditions are oriented around music, graphic arts, and dance. Similarly, prevention activities that involve the entire family may be very appropriate for cultures that are relatively strongly family centered. Overall, the goal is to develop an effective prevention intervention that is also culturally relevant (Castro, Barrera, & Martinez, 2004).

Tailoring Interventions

The concept of shaping the intervention activities to meet the needs of the target group is generally referred to as "tailoring" (e.g., Bartholomew et al., 2006; Hodges & Videto, 2011). Tailoring intervention messages to the characteristics of participants increases the likelihood that they will view the program as relevant and outcomes will be achieved (Witte, 1995). In some types of interventions, such as computer-based "expert systems," messages and strategies are tailored based on characteristics of the individual that have been measured before the intervention begins, such as stage of change (Prochaska & DiClemente, 1984), beliefs, attitudes, and self-efficacy (e.g., Brug, Oenema & Campbell, 2003; Strecher et al., 2000). Addressing participants' perceptions of risk through the delivery of personalized messages in reports, letters, text messages, or computer-assisted instruction is a potentially strong method in motivating health behavior change (DiClemente, Marinilli, Singh, & Bellino, 2001). Some tailoring approaches involve customizing the content and instructional methods of the intervention based on ongoing assessments of the participants' progress in making behavioral changes (Mullen, Brown, & Smith, 1992). For example, an intervention to promote a reduction in dietary fat intake might begin with a computer-based food frequency assessment, followed by feedback about the assessment results that includes a comparison of the participant's fat intake to the recommended intake and the average intake of his/her peers. As participants begin to try the behavior change strategies that are promoted during the intervention, they may receive feedback on their progress on a regular basis (Bartholomew et al., 2006).

Recently, adaptive interventions have become more common in the prevention field. An adaptive intervention is one in which different dosages of certain program components are assigned across participants and/or within participants over time. The dosage that is implemented varies in response to the intervention needs of participants, and dosages are provided based on decision rules that link characteristics of the participant with specific levels and types of program components (Collins, Murphy, & Bierman, 2004). One example of an adaptive intervention is Fast Track (Conduct Problems Prevention Research Group, 1992), a multiyear, multicomponent intervention designed to prevent conduct disorders in high-risk children. In this intervention, parent training and child social skills training components are delivered to all participants, and the home-based counseling and reading tutoring components are delivered adaptively, depending on levels of parental functioning and children's academic performance, respectively. The design of adaptive interventions requires careful consideration of a number of important conceptual and methodological issues, such as identification of adaptive components, choice of tailoring variables, conceptualization and implementation of decision rules, and application of statistical methods to evaluate intervention effectiveness (Collins et al., 2004).

Intervention Settings

In addition to building the content, methods, and components of an intervention, the design team needs to select a setting in which it will be appropriate and feasible to implement the intervention, both during the testing phase and when it is disseminated should it be proven effective. Prevention interventions are implemented in a wide range of settings, such as clinics, schools, churches, worksites, community-based agencies, and others. A review of the literature may identify important considerations in selecting specific intervention settings, such as organizational characteristics that may moderate intervention effectiveness and strategies for ensuring site cooperation and accessibility (D'Onofrio, 2001). However, while prevention researchers may build an intervention around and begin efficacy testing within a specific setting, it has been argued that the sequence of research on an intervention should address its generalizability to the range of settings in which ultimately it may be adopted and implemented (Glasgow, Lichtenstein, & Marcus, 2003).

Stage 3: Pilot Testing

Pilot testing is a critical component of the development of any type of intervention. Sussman (1991) has suggested that pilot testing be conducted at several points during the intervention building process. First, each program activity that has been drafted should be pilot tested with a small number of individuals or groups whose characteristics closely match those of the intervention target group. The activities may be tested individually or in small sets, and some activities may need to be tested more than once if the results suggest the need for significant redesign. After individual activities have been revised based on the pilot test results, a second round of pilot testing is used to determine how to best combine the program activities that have been designed into logical and coherent sessions, segments, or components. For example, the developer of a school-based prevention program may have designed and pilot tested 20 distinct program activities of varying lengths and formats, which need to be combined into ten sessions that will be implemented during 50-min classroom periods over a period of 2 weeks. Pilot testing would enable the development team to determine how to combine and sequence the activities to create a full program that it is feasible to implement within the time available. Finally, a third round of pilot testing builds on the results of the second round, testing the final draft full program or combination of components.

Pilot test data can help the intervention designers determine workability of the activities from the perspective of the program implementer and the likely receptivity of the activities among the target population. Often, pilot test data consist of qualitative observations conducted by program staff. Sussman (1991) suggests that developers also consider conducting studies with single-group or

quasi-experimental designs to test for effects of individual activities or the whole draft program on potential mediators of change. The type of pilot study best suited for a particular intervention will depend on factors such as the program methods (i.e., individual or group), the program setting (e.g., school, clinic), what the developers aim to learn from the piloting, and the resources and time they have available for piloting.

Case Example: Project ALERT Plus

The development of Project ALERT Plus (Longshore, Ellickson, McCaffrey, & St. Clair, 2007) provides an example of how pilot testing can be used to create program components that are feasible for implementation and acceptable to the target audience. ALERT Plus was designed and tested as a high school booster component to Project ALERT, a middle school-based drug abuse prevention curriculum that has shown to prevent or reduce the use of alcohol, marijuana, and tobacco among 8th graders (Ellickson & Bell, 1990; Ellickson, McCaffrey, Ghosh-Dastidar, & Longshore, 2003). The rationale for the booster component was to determine whether additional prevention lessons, implemented during the critical transition period from middle school to high school, would curb the escalating trajectory of substance use among adolescents during the high school years.

At the beginning of the booster component design phase, a review of the literature on school-based substance abuse prevention activities indicated few evidence-based programs targeting high school-aged youth (Tobler, 1986), with the notable exceptions of Project Towards No Drug Abuse (Sussman, Dent, & Stacy, 2002) and a social network program designed by Eggert, Thompson, Herting, Nicholas, and Dicker (1994), both of which targeted adolescents at relatively high risk for substance abuse based on their poor academic performance, regular substance use, and other risk factors. Thus, we found a limited pool of substance abuse prevention activities for use with students in regular high school settings that we might adapt for use in ALERT Plus. Further, the literature suggested that because older adolescents are more likely than younger adolescents to have experimented with drugs, have various situational opportunities for drug use, and have more diverse social networks and activities (e.g., mixing jobs and school), a program approach different from that used in the middle schools years was warranted (Sussman, 1996). The goal was to develop a set of booster lessons for high school youth that had the same theoretical underpinnings and used the same interactive teaching strategies as the middle school program. However, in light of developmental changes during the high school years, greater emphasis would be placed on strengthening norms against high-risk drug use (e.g., drinking and driving), helping students to cope with risky drug situations, and increasing awareness of the consequences of high-risk use (e.g., binge drinking and use of drugs like methamphetamine and ecstasy).

In collaboration with the study investigators, a team of curriculum writers, including the author, adapted a few activities from existing evidence-based

programs and wrote new activities. The ultimate goal was to produce five 50-min curriculum sessions, each comprising several activities. The team developed two versions of each potential program activity, and in the first phase of pilot testing, each version was tested with at least two different classroom groups. By comparing two versions of an activity among similar student groups on the same day, we were able to maximize the limited time and resources we had available for pilot testing. For example, in one activity students discussed acceptability of alcohol use in specific risky situations (e.g., drinking and driving, drinking before school). In one version of the activity, students completed a survey that measured their level of acceptance of their own use; in the other version, each situation was described out loud, followed by students holding up a card indicating their opinion about their own use. To evaluate the workability of the activities, both versions were pilot tested during health classes in a high school in the Los Angeles metropolitan area. The curriculum development team members alternated playing the role of implementer and observer of the activities. When observing, they made general notes and global ratings on variables such as flow, student enthusiasm, and level of interactivity. In addition, students were administered a brief survey before and immediately after each activity to assess responsiveness and effects on hypothesized mediators. Based on these data, we selected the version of the activity that was perceived by the design team as the most workable and that had the stronger impact on students. In the second phase of pilot testing, five sessions that combined several activities were tested to determine whether the sequencing of activities was logical and the length of each session was appropriate. At the end of the process, we had produced a curriculum that was ready for efficacy testing in a large trial (Ellickson, Miller, & Rohrbach, 2000; Longshore et al., 2007).

In summary, pilot testing of both individual activities and the full intervention is a critical stage in the intervention development process. Pilot testing not only allows the design team to assess the feasibility of implementing the intervention activities with the target group but also provides preliminary data on whether the hypothesized mediators of program effect might be changed through the intervention activities that have been designed. It provides the opportunity for the design team to refine the intervention before it expends time, money, and other resources to test its efficacy and effectiveness.

Conclusion

Design is one of the most important and challenging phases in the five-phase life cycle of prevention interventions, which begins with research on the epidemiology of the health problem and ends with dissemination and large-scale implementation of an effective intervention. This chapter has described three key tasks involved in the design phase: adopting a theoretical foundation for the intervention, building the intervention methods and activities, and pilot testing both the individual components and the whole intervention. The design process leads to a theory-based

intervention that is ready for testing to determine its efficacy and effectiveness, as well as the process by which intervention effects are achieved. Application of a theory-based design process and increased detailing of information about the design process and intervention components are important if we are to continue to make advances in the field of prevention science.

One of the key challenges for prevention science in the future is to develop effective interventions that will both generalize to diverse participants and be suitable for implementation in a broad range of settings. The integration of intervention design and implementation approaches is the new frontier in prevention science (National Research Council and Institute of Medicine, 2009), with the ultimate goal of translation of proven prevention interventions into practice and policy on a large scale and in a sustainable way (Spoth et al., 2013). Numerous challenges will need to be addressed to achieve this goal, such as the need for stronger infrastructure for prevention in many healthcare and human service settings, the failure to place a high priority on translation research at the policy level, and many barriers to the development of successful partnerships among prevention scientists, community leaders, and prevention program providers. Ultimately, by designing innovative interventions that are effective, reach diverse participants, and are successfully implemented in a wide range of settings, we should improve the population-level impact of prevention interventions.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bartholomew, L. K., & Mullen, L. K. (2011). Five roles for using theory and evidence in the design and testing of behavior change interventions. *Journal of Public Health Dentistry, 71*, S20–S33.

Bartholomew, L. K., Parcel, G. S., Kok, G., & Gottlieb, N. H. (2006). *Planning health promotion programs: An intervention mapping approach*. San Francisco: Jossey-Bass.

Botvin, G. J., Baker, E., Dusenbury, L., Tortu, S., & Botvin, E. M. (1990). Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: Results of a three-year study. *Journal of Consulting and Clinical Psychology, 58*, 437–446.

Botvin, G. J., Botvin, E. M., Baker, E., Dusenbury, L., & Goldberg, C. J. (1992). The false consensus effect: Predicting adolescents' tobacco use from normative expectations. *Psychological Reports, 70*, 171–178.

Botvin, G. J., Mihalic, S. F., & Grotpeter, J. K. (1998). Life skills training. In D. S. Elliott (Ed.), *Blueprints for violence prevention* (5th ed.). Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado.

Brug, J., Oenema, A., & Campbell, M. (2003). Past, present, and future of computer-tailored nutrition education. *American Journal of Clinical Nutrition, 77*, 1028S–1034S.

Castro, F. G., Barrera, M., & Martinez, C. R. (2004). The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prevention Science, 5*(1), 41–45.

Catania, J. A., Coates, T. J., Stall, R., Bye, L., Kegeles, S. M., Capell, F., et al. (1991). Changes in condom use among homosexual men in San Francisco. *Health Psychology, 10*, 190–199.

Centers for Disease Control and Prevention. (2007a). *Improving public health practice through translation research (R18)*. Retrieved from http://grants.nih.gov/grants/guide/rfa-files/RFA-CD-07-005.html

Centers for Disease Control and Prevention. (2007b). *Best practices for comprehensive tobacco control programs: 2007*. Atlanta, GA: National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. http://www.cdc.gov/tobacco/stateandcommunity/best_practices/pdfs/2007/BestPractices_Complete.pdf.

Cochrane Collaboration. (2012). *Cochrane reviews*. Retrieved from http://www.cochrane.org/cochrane-reviews

Coie, J. D., Watt, N. F., West, S. G., Hawkins, D. J., Asarnow, J. R., Markman, H. J., et al. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist, 48*, 1013–1022.

Collins, L. M., Murphy, S. A., & Bierman, K. L. (2004). A conceptual framework for adaptive preventive interventions. *Prevention Science, 5*(3), 185–196.

Conduct Problems Prevention Research Group. (1992). A developmental and clinical model for the prevention of conduct disorders: The fast track program. *Development and Psychopathology, 4*, 509–528.

D'Onofrio, C. N. (2001). Pooling information about prior interventions: A new program planning tool. In S. Sussman (Ed.), *Handbook of program development for health behavior research and practice* (pp. 158–203). Thousand Oaks, CA: Sage.

DiClemente, C. C., Marinilli, A. S., Singh, M., & Bellino, L. E. (2001). The role of feedback in the process of health behavior change. *American Journal of Health Behavior, 25*, 217–227.

Safe and Drug Free Schools Program, United States Department of Education. (2001). Retrieved from http://www.ed.gov/offices/OESE/SDFS/programs.html

Eggert, L. L., Thompson, E. A., Herting, J. R., Nicholas, L. J., & Dicker, B. G. (1994). Preventing adolescent drug abuse and high school dropout through an intensive school-based social network development program. *American Journal of Health Promotion, 8*, 202–215.

Ellickson, P. L., & Bell, R. M. (1990). Drug prevention in junior high: A multi-site longitudinal test. *Science, 247*, 1299–1305.

Ellickson, P. L., McCaffrey, D. F., Ghosh-Dastidar, B., & Longshore, D. L. (2003). New inroads in preventing adolescent drug use: Results from a large-scale trial of Project ALERT in middle schools. *American Journal of Public Health, 93*, 1830–1836.

Ellickson, P. L., Miller, L., & Rohrbach, L. A. (2000). *ALERT Plus: A drug prevention program for the high school years*. Santa Monica, CA: RAND Corporation.

Elliott, D. S. (1997). *Blueprints for violence prevention* (Vol. 1–11). Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado.

Gilbert, G. G., Sawyer, R. B., & McNeill, E. B. (2011). *Health education: Creating strategies for school and community health* (3rd ed.). Sudbury, MA: Jones and Bartlett.

Glanz, K., Rimer, B. K., & Viswanath, K. (2008). *Health behavior and health education: Theory, research, and practice* (4th ed.). San Francisco: John Wiley & Sons.

Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health, 93*(8), 1261–1267.

Gordon, R. (1983). An operational classification of disease prevention. *Public Health Reports, 98*, 107–109.

Green, L. W., & Kreuter, M. W. (2005). *Health program planning: An educational and ecological approach* (4th ed.). Columbus, OH: McGraw Hill.

Greenwald, P. G. (1990). *Forward. In Smoking, tobacco, and cancer programs*. Rockville, MD: National Cancer Institute.

Hansen, W. B., & Graham, J. W. (1991). Preventing alcohol, marijuana, and cigarette use among adolescents: Peer pressure training versus establishing conservative norms. *Preventive Medicine, 20*, 414–430.

Hodges, B. C., & Videto, D. M. (2011). *Assessment and planning in health programs* (2nd ed.). Sudbury, MA: Jones & Bartlett Publishers.

Issel, L. M. (2009). *Health program planning and evaluation: A practical and systematic approach for community health* (2nd ed.). Sudbury, MA: Jones & Bartlett Publishers.

Jessor, R., & Jessor, S. L. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth*. New York: Academic.

Keefe, J. W. (1987). *Learning style: Theory and practice*. Reston, VA: National Association of Secondary School Principals.

Kegeles, S. D., Hays, R., & Coates, T. J. (1996). The MPowerment project: A community-level HIV prevention intervention for young gay men. *American Journal of Public Health, 86*, 1129–1136.

Kellogg Foundation, W. K. (2004). *Logic model development guide*. Battle Creek, MI: W. K. Kellogg Foundation.

Kelly, J. A., Murphy, D. A., Sikkema, K. J., McAuliffe, T. L., Roffman, R. A., Solomon, L. J., et al. (1997). Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behavior among homosexual men in US cities. *The Lancet, 350*, 1500–1505.

Knowles, M. (1978). *The adult learner: A neglected species* (2nd ed.). Oxford, England: Gulf Publishing.

Longshore, D., Ellickson, P. L., McCaffrey, D. F., & St. Clair, P. A. (2007). School-based drug prevention among at-risk adolescents: Effects of ALERT Plus. *Health Education and Behavior, 34*(4), 651–668.

Mrazek, P., & Haggerty, R. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press. Committee on Prevention of Mental Disorders, Division of Biobehavioral Sciences and Mental Disorders, Institute of Medicine.

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22*, 103–122.

National Institute on Drug Abuse. (2003). *Preventing drug use among children and adolescents: A research-based guide for parents, educators, and community leaders* (2nd ed.). Bethesda, MD: National Institutes of Health.

National Research Council and Institute of Medicine. (2009). Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities. In M. E. O'Connell, T. Boat, & K. E. Warner (Eds.), *Committee on the prevention of mental disorders and substance abuse among children, youth, and young adults: Research advances and promising interventions. Board on children, youth, and families: Division of behavioral and social sciences and education*. Washington, DC: National Academies Press.

Perry, C. L. (1999). *Creating health behavior change: How to develop community-wide programs for youth*. Thousand Oaks, CA: Sage.

Prochaska, J. O., & DiClemente, C. C. (1984). *The transtheoretical approach: Crossing traditional boundaries of therapy*. Homewood, IL: Dow Jones-Irwin.

Rabin, B. A., Brownson, R. C., Haire-Joshu, D., Kreuter, M. W., & Weaver, N. L. (2008). A glossary for dissemination and implementation research in health. *Journal of Public Health Management and Practice, 14*(2), 117–123.

Resnikow, K., Baranowski, T., Ahluwalia, J. S., & Braithwaite, R. L. (1999). Cultural sensitivity in public health: Defined and demystified. *Ethnicity and Disease, 9*, 10–21.

Rohrbach, L. A., D'Onofrio, C. N., Backer, T. E., & Montgomery, S. B. (1996). Diffusion of school-based substance abuse prevention programs. *American Behavioral Scientist, 39*(7), 919–934.

Rohrbach, L. A., English, J., Hansen, W. B., & Johnson, C. A. (2000). Development and pilot testing of Project SMART. In S. Sussman (Ed.), *Handbook of program development for health behavior research and practice* (pp. 425–446). Thousand Oaks, CA: Sage.

Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type 2 translation: Transporting prevention interventions from research to real-world settings. *Evaluation and the Health Professions, 29*, 302–333.

Rohrbach, L. A., Howard-Pitney, B., Unger, J. B., Dent, C. W., Howard, K. A., Cruz, T. B., et al. (2002). Independent evaluation of the California Tobacco Control Program: Relationship between program exposure and outcomes, 1996–1998. *American Journal of Public Health, 92* (6), 975–984.

Rosenfeld, P. L. (2002). The potential of transdisciplinary research for sustaining and extending linkages between the health and social sciences. *Social Science and Medicine, 35*(11), 1343–1357.

Sallis, J. F., Owen, N., & Fisher, E. B. (2008). Ecological models of health behavior. In K. Glanz, B. K. Rimer, & K. Viswanath (Eds.), *Health behavior and health education* (4th ed., pp. 462–484). San Francisco: Jossey-Bass.

Scher, L. S., Maynard, R. A., & Stagner, M. (2006). Interventions intended to reduce pregnancy-related outcomes among adolescents. *Campbell Collaboration Systematic Reviews, 12*.

Spoth, R. S., Rohrbach, L. A., Greenberg, M., Robertson, E., Leaf, P., Brown, C. H., et al. (2013). Addressing core challenges for the next generation of type 2 translation research and systems: The Translation Science to Population Impact (TSci2PI) Framework. *Prevention Science, 14*, 319–351.

Strecher, V. J., Bishop, K. R., Bernhardt, J., Thorp, J. M., Cheuvront, B., & Potts, P. (2000). Quit for keeps: Tailored smoking cessation guides for pregnancy and beyond. *Tobacco Control, 9* (Suppl 3), III78–III79.

Substance Abuse and Mental Health Services Administration. (2007, 2010). *National Registry of Evidence-based Programs and Practices*. Retrieved from http://www.nrepp.samhsa.gov

Sussman, S. (1991). Curriculum development in school-based prevention research. *Health Education Research: Theory and Practice, 6*, 339–351.

Sussman, S. (1996). Development of a school-based drug abuse prevention curriculum for high-risk youths. *Journal of Psychoactive Drugs, 28*(2), 169–182.

Sussman, S., Dent, C. W., Mestel-Rauch, J., Johnson, C. A., Hansen, W. B., & Flay, B. R. (1988). Adolescent nonsmokers, triers, and regular smokers' estimates of cigarette smoking prevalence: When do overestimations occur and by whom? *Journal of Applied Social Psychology, 18*(7), 537–551.

Sussman, S., Dent, C. W., & Stacy, A. (2002). Project towards no drug abuse: A review of the findings and future directions. *America Journal of Health Behavior, 26*, 354–365.

Sussman, S., & Wills, T. A. (2001). Rationale for program development methods. In S. Sussman (Ed.), *Handbook of program development for health behavior research and practice* (pp. 3–30). Thousand Oaks, CA: Sage.

Tobler, N. S. (1986). Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcomes results of program participants compared to a control or comparison group. *Journal of Drug Issues, 15*, 535–567.

Witte, K. (1995). Fishing for success: Using the persuasive health message framework to generate effective campaign messages. In E. Maibach & R. L. Parrott (Eds.), *Designing health messages: Approaches from communication theory and public health practice* (pp. 145–166). Thousand Oaks, CA: Sage.

Wolfson, S. (2000). Students' estimates of the prevalence of drug use: Evidence for a false consensus effect. *Psychology of Addictive Behaviors, 14*(3), 295–298.

Worden, J. K., Flynn, B. S., Geller, B. M., Chen, M., Shelton, L. G., Secker-Walker, R. H., et al. (1988). Development of a smoking prevention mass media program using diagnostic and formative research. *Preventive Medicine, 17*, 531–558.

Chapter 13
Implementation Science and the Effective Delivery of Evidence-Based Prevention

Zili Sloboda, Linda Dusenbury, and Hanno Petras

Introduction

This chapter introduces the topic of implementation; its many phases including dissemination, diffusion and adoption, and sustainability; and those factors that have been found to influence progression along this continuum. Application and implications for prevention service delivery are also discussed. The exciting part about writing this chapter is the simultaneous, parallel emergence of two new scientific fields—prevention science and implementation science. This chapter is an attempt to integrate these two evolving fields.

The chapter builds on the chapter by Rohrbach (2013, Chap. 12) on intervention development and is followed by chapters by Hansen (2013, Chap. 15) on the measurement of implementation fidelity and by van der Kreeft et al. (2014, Chap. 14) on cultural adaptation and implications for training.

Over the past two decades there has been a focus in the social and health services arena including in the prevention field, on the delivery of evidence-based practices (EBPs) grounded in rigorous evaluations and research. At the same time there has been a push to promote dissemination and diffusion of EBPs as rapidly as possible, which has given rise to a new field of study called implementation science (Madon, Hofman, Kupfer, & Glass, 2007). The National Institutes of Health defines implementation science as "the study of methods to promote the integration of research findings and evidence into healthcare policy and practice" (http://www.nlm.nih.gov/hsrinfo/implementation_science.html). Eccles et al. (2009) define implementation research as "the scientific study of methods to promote the systematic uptake of clinical research findings and other evidence-based practices into routine practice, and hence to improve the quality (effectiveness, reliability, safety, appropriateness, equity, efficiency) of health care. It includes the study of influences on

H. Petras (✉)
JBS International, 5515 Security Lane, Suite 800, North Bethesda, MD 20852-5007, USA
e-mail: hpetras@jbsinternational.com

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_13,

healthcare professional and organisational behaviour." Implementation science integrates and extends a number of disciplines including operations research, health services research, industrial engineering, and management science (Schackman, 2010), but it also draws from a range of other related disciplines including decision science, epidemiology, statistics, ethics, sociology, anthropology, and economics.

Implementation issues have taken a central role in prevention as more effective prevention strategies have been developed and evaluated (Wandersman et al., 2008). The importance of understanding how prevention strategies found to be effective through rigorous evaluations are delivered in less controlled circumstances and how variations in delivery impact desired outcomes has prompted a number of significant publications on the issue of implementation in prevention journals (e.g., Greenberg et al., 2003; Greenberg, Domitrovich, Graczyk, & Zins, 2004).

Prevention Programming as Innovation

When we talk about EB prevention programming, in many ways we are talking about innovation (i.e., introducing a new program into an existing system and, in some cases, introducing new ways to deliver a program or services). For this reason, we begin with a discussion of the concept of innovation. Innovation as it relates to the implementation of prevention programming is the introduction of new programming whether it includes the operation of new regulations such as keg registration, "carding," new ways of delivering prevention messages and services such as role play, or a totally new program such as an manual-based prevention curriculum or counseling service (Greenberg et al., 2003, 2004).

Although there are no clear criteria for innovation, there is consensus on the fact that innovation has many dimensions that must be considered when implementation plans are put in effect. Cooper (1998) expressed these as technical versus administrative innovation where the technological innovation would involve an idea that would impact output process while the administrative innovation would include changes in "policies, allocation of resources, and other factors associated with the social structure of the organization" (p. 499).

Another distinction Cooper makes is between product and process innovation. Product innovation reflects changes in the products or services that are developed whereas process innovation focuses on how those products and services are made. This perspective is useful as it suggests that some innovations may be more readily embraced than others. For instance, an administrative or policy change may meet some resistance and may require modest compromises while a change in teaching methodology will more likely meet more resistance and require several steps or phases to get the buy-in necessary for change.

The importance of the type and elements of an innovation has been recognized more recently by prevention researchers. For instance, Greenberg et al. (2004) outline the elements of an intervention including content, structure, dosage, timing,

and so forth, whereas Spoth et al. (2013) discuss the infrastructures that are required to support and sustain an intervention.

Conceptualization of the Implementation Process

A review of the literature on implementation focuses on two endpoints: adoption and sustainability. In a recent article, Spoth et al. (2013) present a research agenda for moving EBPs to scale to ensure population-level impact is laid out. The authors develop a model based on a review of the literature that includes four major phases: pre-adoption, adoption, implementation, and sustainability. It is this framework that guides the discussion regarding implementation.

Conceptual Models of Implementation

A great deal of attention has been given to building conceptual models of implementation building on work from a variety of sources (e.g., Feldstein & Glasgow, 2008; Helfrich et al., 2010; Powell et al., 2012). These models inform strategies to take EBPs to practice and to design research to evaluate them. Furthermore, increased attention has been paid to understanding the relationship between delivery or implementation of an EBP and outcomes of the interventions, services, and practices (Fixsen & Blase, 2009). The model presented here (Fig. 13.1) presents the implementation continuum and those factors and elements that influence the progression from one phase to the next. This figure serves as an outline for this chapter.

Dissemination and Diffusion

The aim of dissemination and diffusion is to actively spread "evidence-based interventions to the target audience via determined channels using planned strategies" (Lomas, 1993; MacLean, 1996; Rabin, Brownson, Haire-Joshu, Kreuter, & Weaver, 2008). The pre-adoption phase in most conceptualizations includes communicating the information regarding EBPs to the most appropriate stakeholders (Sloboda & Schildhaus, 2002). This phase is, perhaps, the most important phase in the implementation process as it brings the innovation to the attention of those who can benefit most from its adoption.

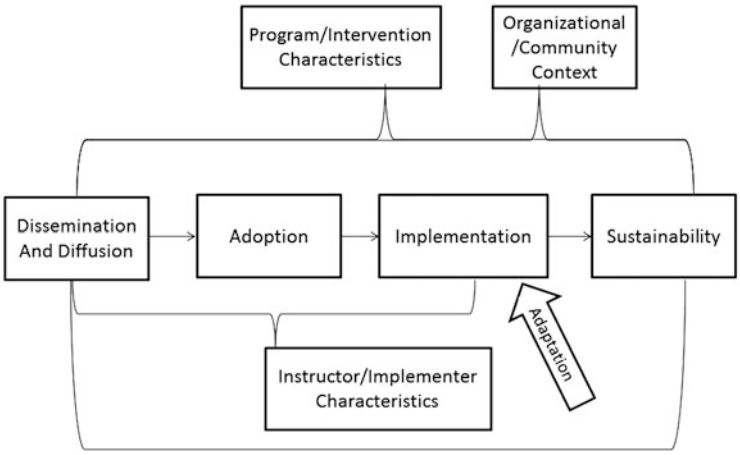

Fig. 13.1 Implementation schema

Dissemination Models

Dissemination can occur from at least three perspectives summarized by Thigpen, Puddy, Singer, and Hall (2012): producer-push, user-pull, and exchange models. The models focus on "who" is active in the process. The first is the producer of knowledge attempting to move research findings to the practitioner community known also as the research-to-practice model (Backer, David, & Soucy, 1995; Klein & Sorra, 1996; Rimer, Glanz, & Rasband, 2001; Rogers, 1995). The second is driven by the consumer of information and can include practitioners or affected individuals or their families (Dawkins et al., 2010; Klein & Sorra, 1996). These individuals seek information or guidance from the best research literature. The third model represents situations where both researchers and practitioners are engaged in sharing knowledge bidirectionally with the practitioners identifying relevant opportunities for research and the researchers working to build practice capacities by using research findings to inform how interventions are delivered (Lomas, 2000).

A number of meta-analyses have been conducted that examine several dissemination producer-push methodologies including printed educational materials (Farmer et al., 2008), guidelines (Grimshaw et al., 2006), educational outreach visits (O'Brien et al., 2007), continuing education meetings and workshops (Forsetlund et al., 2009), and audit and feedback procedures (Ivers et al., 2012). All of these reviews found small effects of the strategy employed and recommend combinations of these (Grimshaw et al., 2006).

User-pull systems have been put in place that present information on research- or evidence-based practices such as the US *Guide to Community Preventive Services* (http://www.thecommunityguide.org/index.html), the *Guide to Clinical Preventive Services* (http://www.ahrq.gov/professionals/clinicians-providers/guidelines-recommendations/guide/), the National Registry of Evidence-Based Programs and

Practices (http://www.nrepp.samhsa.gov/), *European Guidelines on Cardiovascular Disease Prevention*, and others (e.g., Burgers, Grol, Klazinga, Mäkelä, Zaat, & AGREE Collaboration, 2003). Although, no formal evaluations of these systems have been conducted, studies of the adoption of these advocated interventions indicate that only a small percentage of the targeted providers actually incorporate them into practice (e.g., Brownson et al., 2007; Hallfors, Pankratz, & Hartman, 2007; Ringwalt et al., 2002, 2008).

Thigpen et al. (2012) developed the Rapid Synthesis and Translation Process (RSTP) building on the Canadian Health Services Foundation model developed in 1998 that had demonstrated success (Lomas, 2000). Central to this approach was the role served by the foundation as the broker between researchers and policymakers. RSTP was applied to the issue of sexual violence in which research findings are "translated" in usable fashion to policymakers and practitioners in response to their requests. The important elements of this approach are that users of the knowledge are engaged early in conversations with researchers about their own needs and that the products include "actionable knowledge" that can be readily applied to practice. The model, however, has not yet been evaluated.

Other approaches built on the exchange model that have demonstrated success are the Getting to Outcomes (Chinman et al., 2008), Communities That Care (Fagan, Arthur, Hanson, Briney, & Hawkins, 2011), and PROSPER (Crowley, Greenberg, Feinberg, Spoth, & Redmond, 2012). These planning programs include researchers working to provide both guidance and tools necessary to identify and define the problem of concern and the interventions that have the potential to address the problem using and, when necessary, enhancing existing resources.

Social Marketing

Others such as Green, Fettes, and Aarons (2012); Harris et al. (2012); and Wilson, Petticrew, Calnan, and Nazareth (2010) advocate a new conceptual framework for dissemination. Their reviews of existing frameworks underscore the importance of social marketing, persuasion, and communication theories being included in dissemination strategies. Kreuter and Bernhardt (2009) and Dearing, Maibach, and Buller (2006) tie marketing strategies to enhance dissemination. They advocate six action steps for establishing a dissemination infrastructure: promote programs strategically using audience research to identify potential users, to understand their needs, and to create segmented homogeneous groups based on program requirements; build distribution capacity for each user segment; systematically identify all proven programs on an ongoing basis; transform research-tested interventions to make them "ready for use"; build a comprehensive system of user support; and establish evaluation measures and processes to assess the marketing and distribution processes. This view is also strongly supported by Maibach, Van Duyn, and Bloodgood (2006). They point out that marketing is negatively viewed by the public health community as commercial marketing and as supporting

unhealthy eating habits and alcohol and tobacco use. However, they underscore the important role that marketing strategies would have for effective dissemination of evidence-based public health practices.

In a recent systematic review of national quality improvement campaigns, Yuan et al. (2010) found that these campaigns used eight communication strategies: highlighting the evidence base for the practices and the relative simplicity for their delivery; aligning the campaign messages with the strategic goals of the identified adopting organizations; integrating opinion leaders into the enrollment process; forming coalitions of credible sponsors for the campaign; maximizing network exchanges of participating organizations; developing easy-to-use and practical tools and guides; fostering learning opportunities; and incorporating monitoring and evaluation milestones and goals.

It should be emphasized that the type of innovation/intervention that is being disseminated will influence how well received it is and the strategy that is employed. Minor changes in delivery processes, for instance, will be better received than major changes requiring in-depth training. Type of innovation not only is important in the dissemination phase of the implementation process but has implications at all phases. However, this distinction has not been emphasized in the literature.

Adoption/Diffusion

Rogers (2003) called the adoption/diffusion phase the innovation-decision process that leads to adoption, rejection, or discontinuance of the innovation. His conceptualization of diffusion or adoption of an innovation remains a constant in the field. He defines diffusion as "the process through which (1) an innovation (2) is communicated through certain channels (3) over time (4) among the members of a social system" (Rogers, 2002, p. 999). He indicates four main elements related to the diffusion of new ideas including having an innovation, having communication channels in place, time, and the social system. Others have similar definitions; for example, Rabin et al. (2008) state that adoption "is the decision of an organization or community to commit to and initiate an evidence-based intervention."

Rogers examined the rate of adoption of an innovation, that is, the timing between the knowledge about the innovation and the decision to initiate it, and he found that the characteristics of individual adoption of an innovation vary considerably suggesting what has been said above—that dissemination requires multiple strategies to have the broadest reach as possible.

Adoption is characterized by the perception about the relative advantage of the innovation over other ideas/interventions (relative advantage), the extent to which the innovation is perceived as being consistent with existing values, experiences, needs (compatibility), the perceived complexity associated with the implementation of the innovation (complexibility), the perceived degree to which an innovation may be experimented with (trialability), and the perceived degree to which the

results of the innovation are visible to others (observability). Most studies of the adoption of prevention programming use this model or minor variations of it to understand the process and to identify those factors that are associated with adoption. The findings from these studies have formed the basis for planning interventions that attempt to enhance the adoption process.

Barriers and Facilitators to Adoption

Adoption should be viewed as the outcome of a decision-making process. Therefore, who makes the adoption decision, their roles in the organizational structure, and their characteristics are important to the design of dissemination approaches. Furthermore, organizational factors at the state and community levels provide the context for decision making. Curran, Bauer, Mittman, Pyne, and Stetler (2012) add an additional factor, elements of the research itself that may be viewed by decision makers as barriers such as a failure to focus on external validity or implementation challenges and facilitators and real costs of the intervention both initial startup costs and costs to sustain the intervention over time. Most of these factors are also involved in implementation, but they first affect the adoption decision and may well predict the success of implementation later on.

Models of the implementation process acknowledge that the decision-making process to adopt and implement innovative processes and programs needs to address both the personal characteristics and the belief structures of those who will actively deliver the program and of those who are in supervisory positions as well as the organizational climate in which they work (Aarons, 2004, 2005; Aarons & Sommerfeld, 2012; Henggeler et al., 2008; Miller, 2001; Noonan et al., 2009; Parcel et al., 1995; Powers, Bowen, & Bowen, 2010; Rohrbach, Ringwalt, Ennett, & Vincus, 2005; Spell & Blum, 2005).

Rogers summarized strategies for the diffusion of preventive innovations in a presentation he made at an addictions conference (Rogers, 2002). These include the following: change the perceived attributes of prevention innovations stressing their relative advantages, use champions to promote preventive innovations, change the norms of the system regarding preventive innovation through peer support, use entertainment and education to promote innovations by placing educational ideas about prevention in entertainment messages, and activate peer networks to diffuse preventive innovations.

One of the greatest barriers to many prevention efforts is the lack of an established infrastructure to support preventive interventions. Unlike other health and social services, there is usually no central, focused prevention organization. It is perhaps for this reason that when there is local leadership in a community either through individual professionals such as school district-level substance abuse coordinators (Rohrbach et al., 2005), state- and local-level policymakers (Brownson et al., 2007; Hallfors et al., 2007; Jacobs et al., 2012), or community-based partnerships or coalitions, the introduction and implementation of prevention

programming has been successful (Chinman et al., 2008; Crowley et al., 2012; Fagan et al., 2011).

Implementation

Rabin et al. (2008) define implementation as "the process of putting to use or integrating evidence-based interventions within a setting." Fixsen, Naoom, Blase, Friedman, and Wallace (2005), in their review of the literature on implementation issues, suggest that there are six stages of the implementation process. The first already discussed above, adoption, is a stage when an intervention is selected to address a specifically defined need and "fits" within the organizational mandate and operation. The second stage is the process of installing the program. After reviewing the necessary resources and supports needed for the program to be incorporated into an ongoing service delivery system, it is important to put what is needed for the delivery of the program in place: funding, staff training, materials associated with the program/intervention, assigning space, reviewing expected outcomes, monitoring for quality, and so forth. The third phase, initial implementation, determines whether the new program will be fully enmeshed into the organization. It is a period of testing the reality of something new and possibly different being included into the organization and of determining how the program meshes with perceptions of the problem it addresses, the extent to which delivery strategies differ from existing skill levels, and the acceptability of the content and structure of the program. The outcomes from this stage determine whether the program will go to the next stage of implementation, full operation. This fourth stage occurs when the program is integrated into the organizational schedule and becomes "business as usual." Once the program is in place for a time, there may be an additional stage, innovation, when the program may be altered or adapted to meet new client or target population needs or to reflect new staffing visions for the program. The final stage is sustainability—that is, the entrenchment of the program into the organization with associated funding, staffing, and support.

Barriers and Facilitators of Implementation

Aarons, Hurlburt, and Horwitz (2011) indicate two major aspects that impact the success or failure of the implementation of a new program. One dimension is the scale of the implementation efforts: are they small requiring less change in a system or are they large requiring major changes in the organization. The other dimension represents the contextual factors that impact implementation. The outer context would include factors such as funding, the associated sociopolitical situation, the use of contracting to support services that are outside the purview of the organization, interorganizational networks that are sources for information and support for

new services, and the inner contextual factors include such aspects as organizational characteristics, organizational priorities and goals, and readiness for change. Other factors that play a role in implementation success may be related to the nature and stigmatization of the prevention target itself (e.g., sexually transmitted infections and drug abuse) and cultural applicability of the intervention to the target population.

The distinctions made in the literature between the adoption and implementation phases are often quite "fuzzy." The differences may lie primarily at the conceptual level whereas in practice it appears that it is assumed that once an EBP is adopted it is implemented. However, as can be seen from the above discussion, the characteristics of the organization, of those delivering the intervention, or of the target population may dictate the extent to which the intervention is delivered as intended. Green (2008) writes about this issue; the title of his paper "Making research relevant: If it's an evidence-based practice, where's the practice-based evidence?" gets to the heart of the issue. Indeed, how research is funded and how research findings are reported in peer-reviewed journals is not conducive to an easy translation of research into practice. Most researches that appear in peer-reviewed journals and even available on lists of evidence-based prevention programming include primarily efficacy studies, those studies for which the researcher and staff control the training and delivery of the intervention. Replication research or taking these interventions to scale and evaluating them is costly; thus, the generalizability of the findings from efficacy studies to diverse populations is open to question. In addition, often, prevention programming is developed by prevention researchers using specific theories and conceptual models and in many cases developed in a vacuum (i.e., without input from those intended to use them). It is not surprising then that those delivering these programs may want to change them to meet the needs of the provider organization such as a school or community agency and of the target population. There is a fine balance, however, between fidelity to the original content, delivery strategy, and structure of the program and adapting it to meet organizational and target population characteristics and needs. The next section discusses implementation fidelity and adaptation as they relate to high-quality implementation.

High-Quality Implementation

Over the past 15 years, the importance of implementation of a preventive intervention as originally designed has received much attention in the literature. Fixsen et al. (2005) and other implementation scientists have addressed the issue of failure to implement properly as an "implementation gap." This gap can occur at any time, at the time of initiation of a new program or practice or later. The gap can also occur over time with the turnover of staff or, without monitoring or support, as the quality of implementation dissipates (Rohrbach et al., 2005).

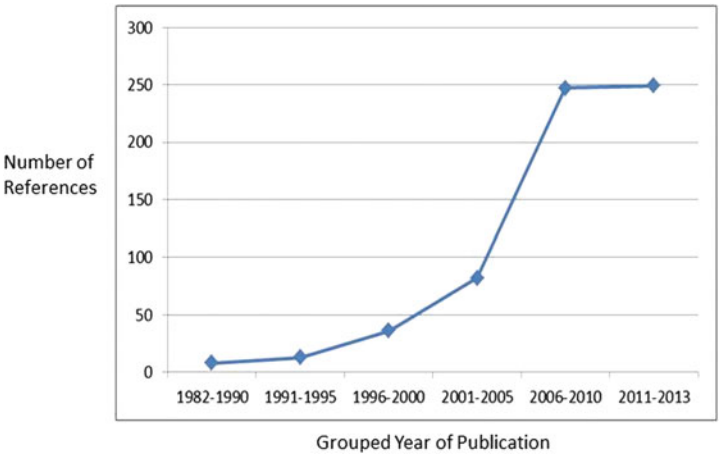

Fig. 13.2 Number of references in PubMed with the term "implementation fidelity" in title or abstract by grouped year of publication

With the greater emphasis on the delivery of evidence-based health, social, and educational services and practices, there have been an increasing number of papers published on implementation fidelity. From 1982 to 1990, only a handful of publications included the term "implementation fidelity" in the title or abstract (i.e., 8), compared with 247 in the period from 2006 to 2010 and 249 from 2011 to January 2013 (Fig. 13.2).

High-quality implementation is associated with better intervention outcomes (e.g., Abbott et al., 1998; Aber, Brown, & Jones, 2003; Battistich, Schaps, Watson, & Solomon, 1996; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Pentz et al., 1990; Resnicow, Cross, & Wynder, 1993; Rohrbach, Graham, & Hansen, 1993; Tobler & Stratton, 1997). This has been demonstrated in research across a number of fields, including education, prevention, and positive youth development. Furthermore, research has revealed that the fidelity of implementation of evidence-based programs is highly variable in regular educational practice (e.g., Botvin, Baker, Dusenbury, Tortu, & Botvin, 1990; Ringwalt et al., 2002; Smith, McCormick, Steckler, & McLeroy, 1993).

Concern about how to ensure high-quality implementation has been addressed in the literature with suggestions related to the improvement of implementation through clarity of the intervention design, manual-based instructions, training, monitoring, and quality improvement (Fixsen et al., 2005). This would include teaching all of the lessons in the appropriate sequence and adhering to or carefully following the instructions in each lesson plan so that all of the points and objectives are covered using the intended methods.

High-quality implementation also has a subjective component; for instance, teachers implement well when they feel they are well prepared, make desired points with clarity, teach with enthusiasm, and create a climate of respect so that students feel safe and engaged in the process. Because every target group is unique in terms

of experiences, needs, and cultural background, implementing well also means knowing when and how a preventive intervention may need to be adapted to achieve its goals, while remaining true to the design of the program—making sure any modifications are aligned with the message and objectives of the program.

Many studies have examined variables that predict or correlate with high-quality implementation. Mihalic (http://www.colorado.edu/cspv/blueprints/Fidelity.pdf) among others have summarized these predictors in five groupings: (1) program characteristics, (2) preplanning involvement, (3) training and technical assistance, (4) organizational integration, and (5) implementer characteristics. For instance, vague or highly complex programs with ambiguous or complicated instructions are more difficult to implement well than clear, simple programs (Payne, 2009; Payne & Eckert, 2010). Documentation of need and the match between this need and the intent of the intervention as well as having implementers involved in the planning process have been found to be associated with high-quality implementation (Gottfredson & Gottfredson, 2002). Studies have confirmed that instructor/implementer training is associated with higher quality implementation (e.g., Ringwalt et al., 2002; Ross, Leupker, Nelson, Saavedra, & Hubbard, 1991; Tappe, Galer-Unti, & Bailey, 1995). Of particular importance is using behavioral teaching strategies in training, making sure the trainees understand the conceptual and theoretical foundation of the intervention and how it works, and availing the implementers of technical assistance as needed (Fagan & Mihalic, 2003).

Organizational climate and community characteristics (e.g., agency stability, shared decision making, chaotic organizational environment, staff turnover) have also been found to relate to quality of implementation (August, Bloomquist, Lee, Realmuto, & Hekter, 2006; Gottfredson & Gottfredson, 2002; Kallestad & Olweus, 2003; Payne, Gottfredon, & Gottferdson, 2006; Rimm-Kaufman, Wanless, Patton, & Deutsch, 2011). In addition, administrative support and implementers' perceptions of support from their administrators for the delivery of a prevention program have been found to predict quality of implementation (Gingiss, Robertson-Gray, & Boerm, 2006; Kam, Greenberg, & Walls, 2003; Payne et al., 2006).

Furthermore, implementer characteristics have been found to predict quality of implementation, including age, teaching experience (Brown, Jones, & Aber, 2011), and prior experience with prevention programming (Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005). Personal characteristics (Domitrovich, Bradshaw, Poduska, Becker, & Ialongo, 2011) such as self-confidence to deliver the program as it was designed, belief in the efficacy of the program (McCormick, Steckler, & McLeroy, 1995; Rohrbach et al., 1993), depression (Hamre & Pianta, 2004), and burnout (Domitrovich et al., 2011; Dusenbury et al., 2005; Ransford, Greenberg, Domitrovich, Small, & Jacobsen, 2009) have been found to be important.

Researchers are increasingly interested in implementation and adaptation in relation to evidence-based programs, and measures and data analysis strategies are becoming more sophisticated to address these issues (e.g., Berkel, Maricio, Schoenfelder, & Sandler, 2011; Hansen, in preparation). Of increasing interest is examining differential outcomes across subgroups to determine whether some groups of individuals are more responsive to an intervention than others (Brown

et al., 2011; Split, Koot, & van Lier, 2013; Supplee, Kelly, MacKinnon, & Barofsky, 2013). For instance, Sloboda et al. (2009) found that high-risk students identified as early marijuana users (use by age 13 years/7th grade) had better outcomes from a substance use prevention program than low-risk, non-substance-using students. It has been suggested that the experience of using substances and related negative consequences of such use make high-risk students more receptive to program messages. Indeed, further mediational analyses indicated that both normative beliefs and refusal skills were associated with nonuse of marijuana by these students when they were in the 11th grade (Teasdale, Stephens, Sloboda, Grey, & Stephens, 2009).

While high-quality implementation has been defined as adhering to the instructions of a program or otherwise implementing activities and lessons as intended by the program developer, the fact is most prevention intervention instructors adapt programs—and they often do so to make the program culturally appropriate for a particular group (e.g., Ringwalt et al., 2008) or because they run out of time (Hill, Maucione, & Hood, 2008). Consistent with these findings, a study of Life Skills Training implementation (Dusenbury et al., 2005) found that all instructors made adaptations.

Backer (2001) defines adaptation as modifications made to an intervention, whether accidental or deliberate. It is possible that some adaptations may enhance program effectiveness, yet there is limited research on how programs are adapted, why they are adapted, and how adaptations affect program outcomes (Castro, Barrera, & Martinez, 2004; Rohrbach, Grana, Sussman, & Valente, 2006). One of the challenges in conducting this type of research is that until recently there were no universal measures of adaptation. For the past several years, Hansen and his research team and other researchers have been working to develop comprehensive measures of adaptation that can be used across classroom-based prevention programs. Hansen (in preparation) conceptualize adaptations as changes in intended methods or messages. In Hansen's system, method adaptations refer to changes in activity structure or instructions, such as adding steps, questions, or examples and stories not intended by the program developers. In contrast, message adaptations refer to new or uncalled for messages about the importance of the program, social norms and attitudes regarding substance use or other risk behaviors, and new concepts or new activities or content designed to promote skill development (e.g., teaching students how to resist peer pressure during a lesson on norm setting). All adaptations, whether method or message, can be either positive or negative and are rated in terms of valence in the Hansen system as ranging from −2 to +2 (Hansen, in preparation).

Enhancing Quality of Implementation

Thus far, the primary strategies for enhancing quality of implementation has been to focus on training, including initial training, ongoing and follow-up training, support

specific to a particular program (for teachers and administrators), coaching, and broader professional development that increases depth of understanding and familiarity with a variety of important teaching methods. Throughout all levels of training, it is important that there be mutual respect and trust so that participants engage and feel safe practicing and receiving feedback (e.g., Ringwalt et al., 2002; Ross et al., 1991; Tappe et al., 1995).

Ideally training, ongoing support, coaching, and broader professional development should accomplish several objectives: (1) engage participants and motivate them to teach the program; (2) provide critically important information that increases depth of understanding about the approach; (3) address misinformation or confusion, as well as resistance; (4) raise important questions (i.e., what does high-quality implementation look like and how do we ensure we are doing it?) to be followed up and developed over time; and (5) introduce a process for launching and sustaining implementation after the initial training through ongoing support, follow-up training, coaching, and professional development.

Evidence-based programs typically offer or require an initial training ranging from one to three days (or more) that provides implementers, such as teachers with an understanding of the framework and rationale for the program, along with training in methods and strategies critical to the program's success. Initial training usually includes some combination of the following elements: (1) important theoretical and research background; (2) program rationale (why this program is important to students and educators); (3) program methods (usually an overview and introduction), identifying those that participants can begin to use right away and relying heavily on a well-organized plan the participant can follow to continue the process of developing techniques and skills for use in the program; (4) an opportunity for investigation and exploration; (5) checks for understanding coupled with an ability on the part of the trainer to correct misinformation and improve depth of understanding; and (6) concluding statement that remind participants of the clearly stated points to know and understand. Further, more comprehensive training, such as ongoing training and coaching, is associated with better outcomes than more limited initial training (Domitrovich et al., 2011; Dusenbury et al., 2010; Rimm-Kaufman et al., 2011; Rohrbach, Gunning, Sun, & Sussman, 2010).

Although training is key, organizational and leadership factors play a significant role in high-quality implementation. Fixsen et al. (2005) provide a summary of these factors indicating the importance of integrating them. They also strongly recommend the development of an implementation team that initiates and supports the provision of new programming into a system. This perspective is advocated by a number of other implementation researchers (e.g., Forman, Olin, Hoagwood, Crowe, & Saka, 2009; Ginexi & Hilton, 2006; Rosenheck, 2001).

Sustainability

Rabin et al. (2008) offer the following definition for sustainability: the extent to which "an evidence-based intervention can deliver its intended benefits over an extended period of time after external support from the donor agency is terminated." They suggest that there are three operational indicators of sustainability: maintenance of initial program benefits, institutionalization of the program in a setting or community (i.e., the extent to which the program is integrated within the culture of the setting/community through policies or practice), capacity building in the recipient setting or community, and/or the extent to which training and durable resources are provided to continue the support of the program.

Wiltsey Stirman et al. (2012) conducted a literature review to explore how sustainability has been defined and to understand how the factors that research had found were associated with sustained programs in healthcare settings. They found that in most cases although the term "sustainability" was used in the reviewed article, the research and article focused on the period after initial implementation. Further, the methodologies used in these studies varied, with some studies being naturalistic or retrospective and some using self-report information rather than independent observations.

Attempts to address sustainability in planning models have been developed, such as PRISM (practical, robust implementation, and sustainability model) (Feldstein & Glasgow, 2008) and ISF (Interactive Systems Framework for Dissemination and Implementation) (Wandersman et al., 2008). These models emphasize the importance of planning for sustainability before program implementation, promoting a positive and supportive climate for the program, and integrating the program into the existing organization and community service system with associated funding (Holder & Moore, 2000; Shediac-Rizkallah, & Bone, 1998).

Integrating EBPs into communities has led to a renewed interest in community partnerships and coalitions, along with the development of tools to help community groups conduct needs assessments, match interventions to need, and conduct ongoing evaluations to monitor progress (Chinman et al., 2008; Hawkins et al., 2008). These efforts have resulted in more effective delivery of EBPs that have been sustained over time (Brown, Feinberg, & Greenberg, 2010; Fagan et al., 2011).

There can be great variability in communities in terms of the availability and strength of leadership, community infrastructure, service delivery systems, and commitment to EBPs, and the readiness of an organization or community to deliver prevention programming should be assessed to determine what groundwork needs to be established before any planning effort (Chilenski, Greenberg, & Feinberg, 2007; Dreisinger et al., 2012; Green et al., 2012; Jacobs et al., 2012; Maclellan-Wright et al., 2007; Stamatakis et al., 2012). These authors point out that organizational and community readiness is key for implementation of prevention processes and programming. Goodman et al. (1998) found that organizations or communities that were rated as having high levels of readiness and that had already

embraced change were more likely to support new approaches. Similarly organizations rated as having low levels of readiness were more likely to resist change.

The organizational literature also emphasizes the importance of readiness as a precursor to implementing change (Armenakis, Harris, & Mossholder, 1993). Entities (e.g., organizations, individuals, or communities) with high levels of readiness are more likely to support a change effort (e.g., Goodman et al., 1998) and sustain new programming. Communities and organizations with low levels of readiness would more likely resist change (e.g., Macrì, Tagliaventi, & Bertolotti, 2002). By focusing on preexisting capacity, levels of readiness can predict behavior (i.e., community support or resistance to change), which then determines the success of the change effort. For this reason, the measurement of community readiness should occur before project implementation (Chilenski et al., 2007).

Conclusion

Two new fields are evolving simultaneously—prevention science and implementation science (e.g., Spoth et al., 2013). In this chapter, we have addressed a number of questions prevention scientists—both researchers and practitioners—will face as they introduce prevention programming into appropriate service organizations and communities. Have we identified the most appropriate research methodologies to study implementation (Curran et al., 2012; Essock, Drake, Frank, & McGuire, 2003; Proctor et al., 2011; Spoth et al., 2013)? What can we learn from the recent research emerging on adaptation, implementation fidelity, and replication? Can prevention program developers recommend adaptations so that interventions achieve desired effects, even when they are altered to meet the needs of a target group (Lipsey, 2003)? How can our understanding of implementation and adaptation inform training and professional development to enhance implementation?

There is a need for replication in research to advance prevention science—for moving evidence-based prevention into practice. Replication is fundamental in all branches of science and is an essential element in the area of implementation science and research (Jasny, Chin, Chong, & Vignieri, 2011; Valentine et al., 2011). To date the discussions on replication do not appear in the implementation literature, although this issue has implications for adaptation and implementation quality.

As others (e.g., Greenberg et al., 2004) have noted, many unknowns and much work remain: first to understand what matters in implementation and adaptation and second to know how to enhance its quality to achieve prevention objectives. The field has made tremendous strides in answering important questions, but much more study of implementation is necessary.

References

Aarons, G. A. (2004). Mental health provider attitudes toward adoption of evidence-based practice: The Evidence-Based Practice Attitude Scale (EBPAS). *Mental Health Services Research, 6*(2), 61–74.

Aarons, G. A. (2005). Measuring provider attitudes toward evidence-based practice: Consideration of organizational context and individual differences. *Child and Adolescent Psychiatric Clinics of North America, 14*(2), 255–271.

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-base practice implementation in public service sectors. *Administration and Policy in Mental Health, 38*, 4–23.

Aarons, G. A., & Sommerfeld, D. H. (2012). Leadership, innovation climate, and attitudes toward evidence-based practice during a statewide implementation. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(4), 423–431.

Abbott, R. D., O'Donnell, J., Hawkins, J. D., Hill, K. G., Kosterman, R., & Catalano, R. F. (1998). Changing teaching practices to promote achievement and bonding to school. *American Journal of Orthopsychiatry, 68*, 542–552.

Aber, J. L., Brown, J. L., & Jones, S. M. (2003). Developmental trajectories toward violence in middle childhood: Course, demographic differences, and response to school-based intervention. *Developmental Psychology, 39*(2), 324–348.

Armenakis, A. A., Harris, S. G., & Mossholder, K. W. (1993). Creating readiness for organizational change. *Human Relations, 46*, 661–703.

August, G. J., Bloomquist, M. L., Lee, S. S., Realmuto, G. M., & Hekter, J. M. (2006). Can evidence-based programs be sustained in community practice settings? The Early Risers' advanced state efficacy trial. *Prevention Science, 7*, 151–165.

Backer, T. E. (2001). *Finding the balance: Program fidelity and adaptation in substance abuse prevention*. Rockville, MD: Center for Substance Abuse Prevention, Substance Abuse and Mental Health Services Administration.

Backer, T. E., David, S. L., & Soucy, G. (1995). Reviewing the behavioral science knowledge base on technology transfer: Introduction. *National Institute on Drug Abuse Monograph, 155*, 1–20.

Battistich, V., Schaps, E., Watson, M., & Solomon, D. (1996). Prevention effects of the Child Development Project: Early findings from an ongoing multisite demonstration trial. Special issue: Preventing adolescent substance abuse. *Journal of Adolescent Research, 11*, 12–35.

Berkel, C., Maricio, A., Schoenfelder, E., & Sandler, I. N. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science, 12*, 23–33.

Botvin, G. J., Baker, E., Dusenbury, L., Tortu, S., & Botvin, E. (1990). Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: Result of a three-year study. *Journal of Consulting and Clinical Psychology, 38*(4), 437–446.

Brown, L. D., Feinberg, M. E., & Greenberg, M. T. (2010). Determinants of community coalition ability to support evidence-based programs. *Prevention Science, 11*(3), 287–297.

Brown, J. L., Jones, S. M., & Aber, J. L. (2011). *Variation in Implementation of School-based Social-Emotional Learning and its Influence on Teacher Development*. Paper presented at the Society for Research in Child Development, Montreal, Canada, April 2

Brown, C. H., Sloboda, Z., Faggiano, F., Teasdale, B., Keller, F., Burkhart, G., et al. (2011). Methods for synthesizing findings on moderation effects across multiple randomized trials. *Prevention Science, 14*, 144–156.

Brownson, R. C., Ballew, P., Dieffenderfer, B., Haire-Joshu, D., Heath, G. W., Kreuter, M. W., et al. (2007). Evidence-based interventions to promote physical activity: What contributes to dissemination by state health departments. *American Journal of Preventive Medicine, 33* (1 Suppl), S66–S73.

Burgers, J. S., Grol, R., Klazinga, N. S., Mäkelä, M., Zaat, J., & AGREE Collaboration. (2003). Towards evidence-based clinical practice: An international survey of 18 clinical guideline programs. *International Journal for Quality in Health Care, 15*(1), 31–45.

Castro, F., Barrera, M., & Martinez, C. R. (2004). The cultural adaptation of prevention interventions: Resolving tension between fidelity and fit. *Prevention Science, 5*(1), 41–45.

Chilenski, S. M., Greenberg, M. T., & Feinberg, M. E. (2007). Community readiness as a multidimensional construct. *Journal of Community Psychology, 35*(3), 347–365.

Chinman, M., Hunter, S. B., Ebener, P., Paddock, S. M., Stillman, L., Imm, P., et al. (2008). The Getting to Outcomes demonstration and evaluation: An illustration of the prevention support system. *American Journal of Community Psychology, 41*, 206–224.

Cooper, J. R. (1998). A multidimensional approach to the adoption of innovation. *Management Decision, 36*(8), 493–502.

Crowley, D. M., Greenberg, M. T., Feinberg, M. E., Spoth, R. L., & Redmond, C. R. (2012). The effect of the PROSPER partnership model on cultivating local stakeholder knowledge of evidence-based programs: A five-year longitudinal study of 28 communities. *Prevention Science, 13*(1), 96–105.

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical Care, 50*, 217–226.

Dawkins, N., Wethington, H., Kettel Khan, L., Grunbaum, J. A., Robin, L., Pitt Barnes, S., et al. (2010). Applying the systematic screening and assessment method to childhood obesity prevention. *New Directions for Evaluation, 125*, 33–49.

Dearing, J. W., Maibach, E. W., & Buller, D. B. (2006). A convergent diffusion and social marketing approach for disseminating proven approaches to physical activity promotion. *American Journal of Preventive Medicine, 31*(4 Suppl), S11–S23.

Domitrovich, C., Bradshaw, C., Poduska, J., Becker, K., & Ialongo, N. (2011). *Enhancing the Fidelity of a Classroom-based, Social-Emotional Intervention through Coaching*. Paper presented at the Society for Research in Child Development, Montreal Canada

Dreisinger, M. L., Boland, E. M., Filler, C. D., Baker, E. A., Hessel, A. S., & Brownson, R. C. (2012). Contextual factors influencing readiness for dissemination of obesity prevention programs and policies. *Health Education Research, 27*(2), 292–306.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school based universal interventions. *Child Development, 82*, 405–432.

Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., & Falco, M. (2005). Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research, 20*, 308–313.

Dusenbury, L., Hansen, W. B., Jackson-Newsom, J., Pittman, D., Wilson, C., Simley, K., et al. (2010). Coaching to enhance quality of implementation in prevention. *Health Education, 110*, 43–60.

Eccles, M. P., Armstrong, D., Baker, R., Cleary, K., Davies, H., Davies, S., et al. (2009). An implementation research agenda. *Implementation Science, 4*, 18.

Essock, S. M., Drake, R. E., Frank, R. G., & McGuire, T. G. (2003). Randomized controlled trials in evidence-based mental health care: Getting the right answer to the right question. *Schizophrenia Bulletin, 29*(1), 115–123.

Fagan, A. A., Arthur, M. W., Hanson, K., Briney, J. S., & Hawkins, J. D. (2011). Effects of Communities That Care on the adoption and implementation fidelity of evidence-based prevention programs in communities: Results from a randomized controlled trial. *Prevention Science, 12*, 223–234.

Fagan, A. A., & Mihalic, S. (2003). Strategies for enhancing the adoption of school-based prevention programs: Lessons learned from the Blueprints for Violence Prevention replications of the Life Skills Training program. *Journal of Community Psychology, 31*(3), 235–253.

Farmer, A. P., Légaré, F., Turcot, L., Grimshaw, J., Harvey, E., McGowan, J. L., et al. (2008). Printed educational materials: Effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews, 2008*, CD004398.

Feldstein, A. C., & Glasgow, R. E. (2008). A Practical, Robust Implementation and Sustainability Model (PRISM) for integrating research findings into practice. *Joint Commission Journal on Quality and Patient Safety, 34*(4), 228–243.

Fixsen, D., & Blase, K. (2009). Technical assistance in special education: Past, present, and future. *Topics in Early Childhood Special Education, 29*(1), 62–64.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: National Implementation Research Network. Retrieved from http://www.fpg.unc.edu/~nirn/resources/.publications/Monograph/pdf/Monograph_full.pdf.

Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health, 1*, 26–36.

Forsetlund, L., Bjørndal, A., Rashidian, A., Jamtvedt, G., O'Brien, M. A., Wolf, F., et al. (2009). Continuing education meetings and workshops: Effects on professional practice and health care outcomes. *Cochrane Database Systematic Review, 2*, CD003030.

Ginexi, E., & Hilton, T. (2006). What's next for translation research? *Evaluation and the Health Professions, 29*(3), 334–347.

Gingiss, P. M., Robertson-Gray, C., & Boerm, M. (2006). Bridge it: A system for predicting implementation fidelity for school-based tobacco prevention programs. *Prevention Science, 7*, 197–207.

Goodman, R., Speers, M., Mcleroy, K., Fawcett, S., Kegler, M., Parker, E., et al. (1998). Identifying and defining the dimensions of community capacity to provide a base for measurement. *Health Education and Behavior, 25*(3), 258–278.

Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency, 39*, 3–35.

Green, L. W. (2008). Making research relevant: If it is an evidence-based practice, where's the practice-based evidence? *Family Practice, 25*(Suppl 1), i20–i24.

Green, A. E., Fettes, D. L., & Aarons, G. A. (2012). A concept mapping approach to guide and understand dissemination and implementation. *Journal of Behavioral and Health Services Research, 39*, 362–373.

Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2004). *The study of implementation in school-based preventive interventions: Theory, research, and practice*. Washington, DC: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*(6–7), 466–474.

Grimshaw, J., Eccles, M., Thomas, R., MacLennan, G., Ramsay, C., Fraser, C., et al. (2006). Toward evidence-based quality improvement: Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies, 1966–1998. *Journal of General Internal Medicine, 21*, S14–S20.

Hallfors, D. D., Pankratz, M., & Hartman, S. (2007). Does federal policy support the use of scientific evidence in school-based prevention programs? *Prevention Science, 8*(1), 75–81.

Hamre, B., & Pianta, R. (2004). Self-reported depression in nonfamilial caregivers: Prevalence and associations with caregiving behaviors in child care settings. *Early Childhood Research Quarterly, 19*, 297–318.

Hansen, W. B. (in preparation), *Styles of adaptation: How teachers modify All Stars and how styles of chance impact fidelity and prevention*

Harris, J. R., Cheadle, A., Hannon, P. A., Forehand, M., Lichiello, P., Mahoney, E., et al. (2012). A framework for disseminating evidence-based health promotion practices. *Preventing Chronic Disease, 9*, 111001. doi:10.5888/pcd9.110081.

Hawkins, J. D., Catalano, R. W., Arthur, M. W., Brown, E. C., Abbott, R. D., & Murray, D. M. (2008). Testing communities that care: The rationale, design and behavioral baseline equivalence of the community youth development study. *Prevention Science, 9*, 178–190.

Helfrich, C. D., Damschroder, L. J., Hagedorn, H. J., Daggett, G. S., Sahay, A., Ritchie, M., et al. (2010). A critical synthesis of literature on the promoting action on research implementation in health services (PARIHS) framework. *Implementation Science, 5*, 5–82.

Henggeler, S. W., Chapman, J. E., Rowland, M. D., Halliday-Boykins, C. A., Randall, J., et al. (2008). Statewide adoption and initial implementation of contingency management for substance-abusing adolescents. *Journal of Consulting and Clinical Psychology, 76*(4), 556–567.

Hill, L. G., Maucione, K., & Hood, B. K. (2008). A focused approach to assessing program fidelity. *Prevention Science, 8*, 25–34.

Holder, H. D., & Moore, R. S. (2000). Institutionalization of community action projects to reduce alcohol-use related problems: Systematic facilitators. *Substance Use and Misuse, 35*(1–2), 75–86.

Ivers, N., Jamtvedt, G., Flottorp, S., Young, J. M., Odgaard-Jensen, J., French, S. D., et al. (2012). Audit and feedback: Effects on professional practice and healthcare outcomes. *Cochrane Database of Systematic Reviews, 6*, CD000259.

Jacobs, J. A., Clayton, P. F., Dove, C., Funchess, T., Jones, E., Perveen, G., et al. (2012). A survey tool for measuring evidence-based decision making capacity in public health agencies. *BMC Health Services Research, 9*, 12–57.

Jasny, B. R., Chin, G., Chong, L., & Vignieri, S. (2011). Introduction again, and again, and again. *Science, 2*, 1225.

Kallestad, J. H., & Olweus, D. (2003). Predicting teachers and schools implementation of the Olweus Bullying Prevention Program. *Prevention and Treatment, 6*(1), 3–21.

Kam, C., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science, 4*, 55–63.

Klein, K. J., & Sorra, J. S. (1996). The challenge of innovation implementation. *The Academy of Management Review, 2*, 1055–1080.

Kreuter, M. W., & Bernhardt, J. M. (2009). Reframing the dissemination challenge: A marketing and distribution perspective. *American Journal of Public Health, 99*(12), 2123–2127.

Lipsey, M. W. (2003). The challenges of interpreting research for use by practitioners comments on the latest products from the task force on community preventive services. *American Journal of Preventive Medicine, 28*(2), S1–S3.

Lomas, J. (1993). Diffusion, dissemination, and implementation: Who should do what? *Annals of the New York Academy of Science, 703*, 226–235.

Lomas, J. (2000). Using "linkage and exchange" to move research into policy at a Canadian foundation. *Health Affairs, 19*, 236–240.

MacLean, D. R. (1996). Positioning dissemination in public health policy. *Canadian Journal of Public Health, 87*(Suppl 2), S40–S43.

Maclellan-Wright, M. F., Anderson, D., Barber, S., Smith, N., Cantin, B., Felix, R., et al. (2007). The development of measures of community capacity for community-based funding programs in Canada. *Health Promotion International, 22*(4), 299–306.

Macrì, D. M., Tagliaventi, M. R., & Bertolotti, F. (2002). A grounded theory for resistance to change in a small organization. *Journal of Organizational Change Management, 15*(3), 292–310.

Madon, T., Hofman, K. J., Kupfer, L., & Glass, R. I. (2007). Implementation science. *Science, 318*, 1728–1729.

Maibach, E. W., Van Duyn, M. A. S., & Bloodgood, B. (2006). A marketing perspective on disseminating evidence-based approaches to disease prevention and health promotion. *Preventing Chronic Disease, 3*(3), 1–11.

McCormick, L. K., Steckler, A. B., & McLeroy, K. R. (1995). Diffusion of innovations in schools: A study of adoption and implementation of school-based tobacco prevention curricula. *American Journal of Health Promotion, 9*(3), 210–219.

Miller, R. L. (2001). Innovation in HIV prevention organizational and intervention characteristics affecting program adoption. *American Journal of Community Psychology, 29*(4), 621–647.

Noonan, R. K., Emshoff, J. G., Mooss, A., Armstrong, M., Weinberg, J., & Ball, B. (2009). Adoption, adaptation, and fidelity of implementation of sexual violence prevention programs. *Health Promotion Practice, 10*(1), 59S–70S.

O'Brien, M. A., Rogers, S., Jamtvedt, G., Oxman, A. D., Odgaard-Jensen, J., Kristoffersen, D. T., et al. (2007). Educational outreach visits: Effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews, 4*, CD000409.

Parcel, G. S., O'Hara-Tompkins, N. M., Harrist, R. B., Basen-Engquist, K. M., McCormick, L. K., Gottlieb, N. H., et al. (1995). Diffusion of an effective tobacco prevention program, Part II: Evaluation of the adoption phase. *Health Education Research, 10*(3), 297–307.

Payne, A. (2009). Do predictors of the implementation quality of school-based prevention program differ by program type. *Prevention Science, 10*, 151–167.

Payne, A. A., & Eckert, R. (2010). The relative importance of provider, program, school and community predictors of the implementation quality of school-based prevention programs. *Prevention Science, 11*, 126–141.

Payne, P., Gottfredon, D. C., & Gottferdson, G. D. (2006). School predictors of the intensity of implementation of school-based prevention programs: Results from a national study. *Prevention Science, 7*, 225–237.

Pentz, M. A., Trebow, E. A., Hansen, W. B., MacKinnon, D. P., Dwyer, J. H., Johnson, C. A., et al. (1990). Effects of program implementation on adolescent drug use behavior: The Midwestern Prevention Project. *Evaluation Review, 14*, 264–289.

Powell, B. J., McMillen, J. C., Proctor, E. K., Carpenter, C. R., Griffey, R. T., Bunger, A. C., et al. (2012). A compilation of strategies for implementing clinical innovations in health and mental health. *Medical Care and Review, 69*(2), 123–157.

Powers, J. D., Bowen, N. K., & Bowen, G. L. (2010). Evidence-based programs in school settings: Barriers and recent advances. *Journal of Evidence-Based Social Work, 7*, 313–331.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., et al. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health, 38*(2), 65–76.

Rabin, B. A., Brownson, R. C., Haire-Joshu, D., Kreuter, M. W., & Weaver, N. L. (2008). A glossary for dissemination and implementation research in health. *Journal of Public Health Management and Practice, 14*(2), 117–123.

Ransford, C., Greenberg, M., Domitrovich, C., Small, M., & Jacobsen, L. (2009). The role of teachers' psychological experience and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *School Psychology Review, 38*, 510–532.

Resnicow, K., Cross, D., & Wynder, E. (1993). The know your body program: a review of studies. *Bulletin of the New York Academy of Medicine, 70*, 188–207.

Rimer, B. K., Glanz, K., & Rasband, G. (2001). Searching for evidence about health education and health behavior interventions. *Health Education and Behavior, 28*(2), 231–248.

Rimm-Kaufman, S. E., Wanless, S., Patton, C., & Deutsch, N. (2011). *Teachers' accounts of the process of teacher change: Examining fidelity of implementation*. Paper presented at the Society for Research in Child Development, Montreal Canada

Ringwalt, C. L., Ennett, S., Vincus, A., Thorne, J., Rohrbach, L. A., & Simons-Rudolph, A. (2002). The prevalence of effective substance use prevention curricula in U.S. middle schools. *Prevention Science, 3*, 257–265.

Ringwalt, C., Hanley, S., Vincus, A. A., Ennett, S. T., Rohrbach, L. A., & Bowling, J. M. (2008). The prevalence of effective substance use prevention curricula in the nation's high schools. *Journal of Primary Prevention, 29*, 479–488.

Rogers, E. M. (1995). Lessons for guidelines from the diffusion of innovations. *Joint Commission Journal on Quality Improvement, 21*(7), 324–328.

Rogers, E. M. (2002). Diffusion of preventive innovations. *Addictive Behaviors, 27*, 989–993.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine, 22*, 237–260.

Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type II translation: Transporting prevention interventions from research to real-world settings. *Evaluation for Health Professionals, 29*, 302.

Rohrbach, L. A., Gunning, M., Sun, P., & Sussman, S. (2010). The project towards no drug abuse (TND) dissemination trial: implementation fidelity and immediate outcomes. *Prevention Science, 11*, 77–88.

Rohrbach, L. A., Ringwalt, C. L., Ennett, S. T., & Vincus, A. A. (2005). Factors associated with adoption of evidence-based substance use prevention curricula in the U.S. school districts. *Health Education Research, 20*(5), 514–526.

Rosenheck, R. (2001). Stages in the implementation of innovative clinical programs in complex organizations. *Journal of Nervous and Mental Disease, 189*(12), 812–821.

Ross, J. G., Leupker, R. V., Nelson, G. D., Saavedra, P., & Hubbard, B. (1991). Teenage health teaching modules: Impact of teacher training on implementation and student outcomes. *Journal of School Health, 61*, 31–34.

Schackman, B. R. (2010). Implementation science for the prevention and treatment of HIV/AIDS. *Journal of Acquired Immune Deficiency Syndrome, 55*(Suppl 1), S27–S31.

Shediac-Rizkallah, M. C., & Bone, L. R. (1998). Planning for the sustainability of community-based health programs: Conceptual frameworks and future directions for research, practice and policy. *Health Education Research, 13*(1), 87–108.

Sloboda, Z., & Schildhaus, S. (2002). A discussion of the concept of technology transfer of research-based drug "abuse" prevention and treatment interventions. *Substance Use and Misuse, 37*(8–10), 1079–1087.

Sloboda, Z., Stephens, R. C., Stephens, P. C., Grey, S. F., Teasdale, B., Hawthorne, R. D., et al. (2009). The Adolescent Substance Abuse Prevention Study: A randomized field trial of a universal substance abuse prevention program. *Drug and Alcohol Dependence, 102*, 1–10.

Smith, D. W., McCormick, L. K., Steckler, A. B., & McLeroy, K. R. (1993). Teachers use of health curricula: Implementation of Growing healthy, Project SMART and Teenage Health Teaching Modules. *Journal of school Health, 63*, 349–354.

Spell, C. S., & Blum, T. C. (2005). Adoption of workplace substance abuse prevention programs: Strategic choice and institutional perspectives. *Academy of Management Journal, 48*(6), 1125–1142.

Split, J. L., Koot, J. M., & van Lier, P. A. C. (2013). For whom does it work? Subgroup differences in the effects of a school-based universal prevention program. *Prevention Science, 14*, 479–488.

Spoth, R., Rohrbach, L. A., Greenberg, M., Leaf, P., Brown, H., Fagan, A., et al. (2013). Addressing challenges for the next generation of type 2 translation research: The translation science to population impact framework. *Prevention Science, 14*, 319–351.

Stamatakis, K. A., McQueen, A., Filler, C., Boland, E., Dreisinger, M., Brownson, R. C., et al. (2012). Measurement properties of a novel survey to assess stages of organizational readiness for evidence-based interventions in community chronic disease prevention settings. *Implementation Science, 7*, 65.

Supplee, L. H., Kelly, B. C., MacKinnon, D. M., & Barofsky, M. Y. (2013). Introduction to the special issue: Subgroup analysis in prevention and intervention research. *Prevention Science, 14*, 107–110.

Tappe, M. K., Galer-Unti, R. A., & Bailey, K. C. (1995). Long term implementation of Teenage health Teaching Modules by trained teachers: A case study. *Journal of School Health, 65*(10), 411–415.

Teasdale, B., Stephens, P. C., Sloboda, Z., Grey, S. F., & Stephens, R. C. (2009). The influence of program mediators on outcomes for substance users and non-users at baseline. *Drug and Alcohol Dependence, 102*, 11–18.

Thigpen, S., Puddy, R. W., Singer, H. H., & Hall, D. M. (2012). Moving knowledge into action: Developing the rapid synthesis and translation process within the interactive systems framework. *American Journal of Community Psychology, 50*(3–4), 285–294.

Tobler, N. S., & Stratton, H. H. (1997). Effectiveness of school-based drug prevention programs: A meta-analysis of the research. *Journal of Primary Prevention, 18*, 71–128.

Valentine, J. C., Biglan, A., Boruch, R. F., Castron, F. G., Collins, L. M., Flay, B. R., et al. (2011). Replication in prevention science. *Prevention Science, 12*, 102–117.

van der Kreeft, P., Jongbloet, J., & Van Havere, T. (2014). Factors impacting implementation: Cultural adaptation and training. In Z. Sloboda & H. Petras (Eds.), *Advances in prevention science: Defining prevention science (chapter 14)*. New York: Springer.

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, R., Stillman, L., et al. (2008). Bridging the gap between prevention research and practice: The Interactive Systems Framework for dissemination and implementation. *American Journal of Community Psychology, 41*, 171–181.

Wilson, P. M., Petticrew, M., Calnan, M. W., & Nazareth, I. (2010). Disseminating research findings: What should researchers do? A systematic scoping review of conceptual frameworks. *Implementation Science, 5*, 91.

Wiltsey Stirman, S., Kimberly, J., Cook, N., Calloway, A., Castro, F., & Charns, M. (2012). The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. *Implementation Science, 7*, 17.

Yuan, C. T., Nembhard, I. M., Stern, A. F., Brush, J. E., Jr., Krumholz, H. M., & Bradley, E. H. (2010). Blueprint for the dissemination of evidence-based practices in health care. *Commonwealth Fund Publication, 86*, 1–14. Retrieved from http://www.commonwealthfund.org/~/media/Files/Publications/Issue%20Brief/2010/May/1399_Bradley_blueprint_dissemination_evidencebased_practices_ib.pdf.

Chapter 14
Factors Affecting Implementation: Cultural Adaptation and Training

Peer van der Kreeft, Johan Jongbloet, and Tina Van Havere

Introduction

If a school wishes to set up a new prevention program, it is hoped it makes choices based on evidence and not on opinion. To implement an evidence-based intervention, the ideal process is to search again for best practice instructions on how to do so, but such instructions are often vague or not supported by research. Many times, the school follows a general rule: "stick to exact application of the program as designed." That, however, does not seem to be possible in the classroom. This chapter contributes to the discourse on how to improve instructions for quality program implementation in school-based prevention programs and can be an inspiration for implementation in other settings, too.

Accurate and high-fidelity implementation of prevention is important to examine whether a program works and under which circumstances it does so. Implementation fidelity can make a difference between a positive or a negative outcome of a program. Fidelity, sometimes referred to as adherence, is the extent to which specified program components were delivered as prescribed and is the most commonly measured dimension of implementation (Durlak & DuPre, 2008). Not only fidelity but also cultural adaptation might be needed for an intervention to be appropriate for a specific region or a particular target population, and specific adaptations could even be required at the level of the school. Moreover, research suggests that when target populations or cultural settings are different from those in the research trial, adaptations to the intervention do not necessarily make the program less effective. On the contrary, it could even prove to be more effective than it was in the original research (Kelly et al., 2000). What adaptations do we have in view if we try to balance fidelity to the original with fit to the target group? One group of adaptations concerns content: the risk factors and protective factors.

P. van der Kreeft (✉)
Department of Education, Health and Social Work, University College Ghent, Voskenslaan 362, Ghent BE9000, Belgium
e-mail: peer.vanderkreeft@hogent.be

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_14,

Violent or unhealthy behavior may be caused by copying peers' behavior in one country and by parental examples in another. It would be highly ineffective to adopt a story depicting "look at what your parents do" without any adaptation of the content. The given example is a difference in the "deep structure" of the intervention, which is more difficult to identify than the "surface structure", which includes trivial differences between countries, regions, or otherwise differentiated cultures (e.g., sports, food). Identifying deep structure differences requires analysis by an expert team and often a pilot of the already evidence-based intervention in a new context. A second group of adaptations is formed by the instructions, logistics, and formative components of the intervention. It is obvious that, when sophisticated training of local prevention workers is already provided, capacity building will be less necessary than when we start from scratch. This also affects the thoroughness of manuals and handbooks: literally translating a program that has been tested in a region where deliverers need a lot of rudimentary information may cause resistance in a region where the program implementers already have received further training, thus decreasing the proven effectiveness potential. Program instructions, distinguished from program content, also encompass the directives for exercises, time group management, and group dynamics. In these areas it also is necessary to adapt to national, regional, or local circumstances. Program designers cannot, however, endlessly differentiate between potential groups that want to implement the intervention; that is a task for the implementing team. However, teachers, as one example of program deliverers, tend to bear the burden of many of the necessary tasks, which does not allow them the time to make their own tailor-made implementation plans. Therefore, developers and designers of prevention interventions should facilitate adaptation or create a built-in adaptation protocol.

Stages in Effective Drug Prevention Programming

Because of its proven effectiveness in changing behavior, adopting a Comprehensive Social Influence (CSI) program as a national or regional standard or as a curriculum requirement grants the responsible authority a reliable indicator of how prevention is implemented at the school level. Such a program provides quality control and, in the case of "Unplugged" from the European Drug Abuse Prevention Trial (EU-Dap), ensures that the prevention program being implemented is based on evidence (EU-Dap Consortium, 2010).

The chain of actions in Fig. 14.1 shows the distinction of actors from the (inter) national to the class level. Program design, including trial and effectiveness evaluation, often occurs on the national and sometimes international level. Results from the effectiveness evaluation will include monitoring or other process elements, leading to improvement of the original program at the designer's level. Adoption of the evidence-based program is initiated by a national, regional, or local prevention team and can include adaptation of the materials. Adaptation also occurs through the introduction of the program to a school; for example, a local prevention

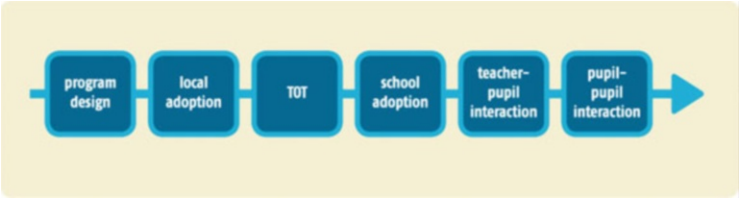

Fig. 14.1 Separate actors in the chain of actions from a Comprehensive Social Influence school-based program design to pupil–pupil interaction

worker might emphasize an issue that is highlighted in local health policy, or the trainer of a teacher workshop may give more attention to a specific protective factor if one of the participants has particular expertise in it. Finally, a teacher adapts program delivery according to an assessment of situational and circumstantial conditions of the class and individual students. The ultimate target of a CSI program – interactivity between pupils –thus will be different at the final stage compared to the program design conditions.

Crucial in this chain of actions is that the intervention actors have at hand a distinction between components that are necessary for the proven effectiveness and others that are subject to adaptation. Expense and burden of the necessary analysis for this are high (Castro, Barrera, & Holleran Steiker, 2011); nevertheless, mediator analysis of an evidence-based program – it has the desired effect, but what exactly makes it effective? – or even monitoring and other process elements may contribute to this distinction. It is a part of implementation that will be actualized in any case, with or without the involvement of the developers. It is better to continue participation in the process.

Impact Factors of Implementation and Adaptation

The quality of intervention delivery includes the methods used by the facilitators, such as interactive learning, but participants' behavior also seems to influence the quality of implementation (Berkel, Mauricio, Schoenfelder, & Sandler, 2011; Haystead & Marzano, 2009). A program design will contain guidelines for the teacher on how to instruct pupils and the prevention worker to address parents. Many components at this level of the intervention still are not translated into written instructions and are not documented. In fact, they are not standardized because in the last stages of implementation a lot must be consigned to the deliverer with his or her specific capacities: teacher to pupil, prevention worker to parent or group. Instruction at this level can be grasped only in the content and process of training.

Adaptation of evidence-based programs will be a growing area of experience and research. Not only has the need for proven effectiveness in prevention responses increased worldwide but also the dissemination of intervention concepts, program materials, and study results over international borders is growing at an unmatched pace. Implementation of the Unplugged intervention in more than 24 countries (described later in this chapter) illustrates the following impact factors for program adaptation:

1. Integrating a relevant share of interactive activities in the program, which increases the degree of participation and contribution of the final target group in intervention components.
2. Designing the original program in a multicultural environment, including countries with explicit differences relevant to the program components.
3. Distinguishing core components from components submitted to adaptation in the program description in a dynamic concept where field experiences and further studies can add up to the initial set.
4. Acquiring and selecting trainers who will have to build the skills of program deliverers to adjust and fine-tune programs to their target groups.
5. Using training of trainers (TOT), supervision, and continued international exchange of trainer's expertise through digital as well as face-to-face encounters to ameliorate possible adaptations.
6. Involving a multidisciplinary adaptation group with expertise in content and process elements.
7. Inserting a limited intervention trial only for adaptation purposes.
8. Scheduling a local effectiveness study after an adequate adaptation period to test the adapted version in the intended target group.

Prevention often is delivered to schools without being called for and tends to be imposed on them. It is therefore essential to ensure that not only effectiveness but also the risks of iatrogenic effects are part of the equation in decision making. Information about this decision-making process and the risks connected to opinion-based instead of evidence-based program choices is part of a sound implementation model.

This chapter focuses on cultural adaptation and training as factors of impact on qualitative implementation of a drug prevention program, illustrated by the international school-based intervention "Unplugged." We will look at social processes at a level of cultural adaptation, as when an evidence-based program moves from one region to another with a different language and culture. Cultural brokerage is a prominent aspect of this process. Second, we will look for possible pitfalls when widely disseminating such prevention programs. Schools themselves will have to adapt and integrate the program into their specific environments. On both levels we emphasize the importance and the process of training trainers and teachers as program deliverers.

Dissemination of the EU-Dap Unplugged Program

In 2002–2005 the EU-Dap project conducted a randomized controlled trial of "Unplugged," a school-based drug prevention program for young people from 12 to 14 years old. It was the first European trial in that area and showed promising effectiveness results. Nine regional centers collaborated on implementing and cross-evaluating the 12-lesson Unplugged program among 7,079 students from

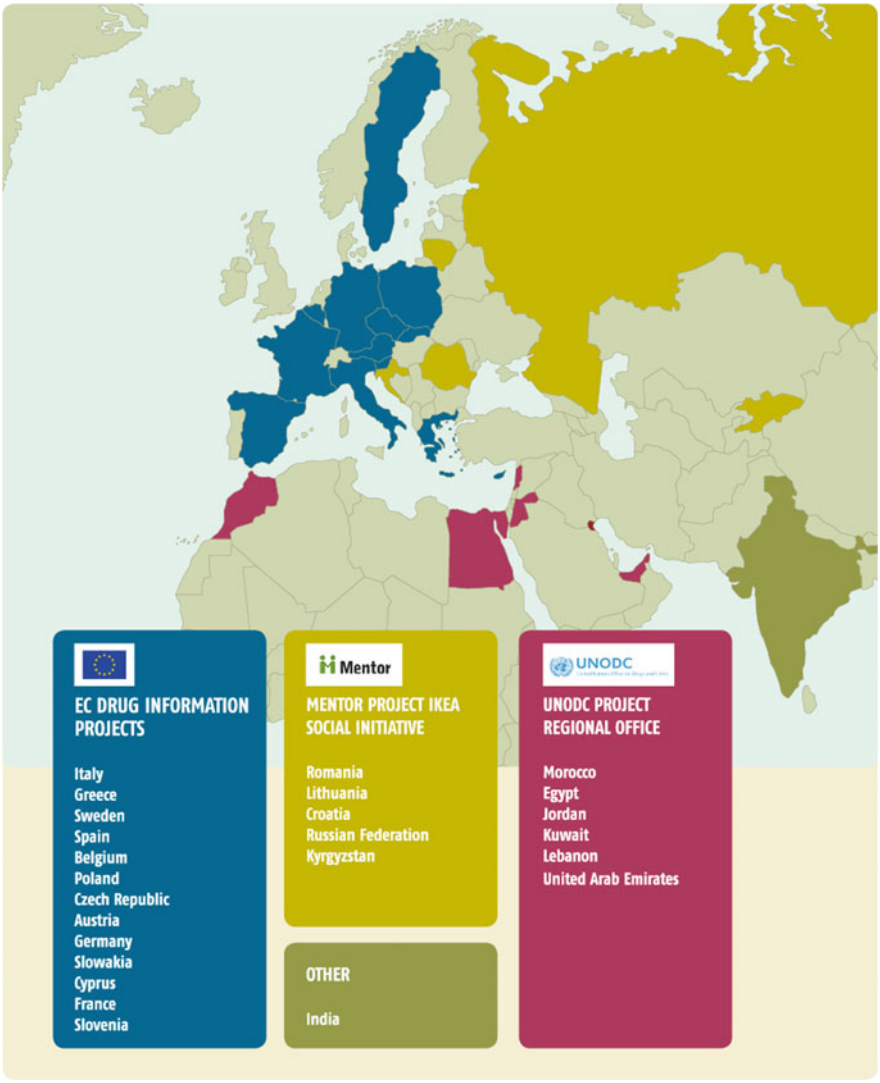

Fig. 14.2 Unplugged international dissemination as of 3 November 2011

143 schools in 7 countries (Italy, Sweden, Austria, Belgium, Spain, Greece, and Germany). From 2006 to 2009 a follow-up project called EU-Dap 2 focused on improvement of the program components, dissemination of the effective program, involving more European Union (EU) member states, and evaluating the effect of an implementation instrument (Faggiano et al., 2010). These two EU Public Health cofunded projects delivered a network of national and regional centers in the EU working with the same standardized and evidence-based drug prevention program "Unplugged" (Van der Kreeft et al., 2009).

Since its design in 2004 Unplugged has evolved as a multiple-adapted program from the EU-Dap network, which was improved, evaluated, and found to be evidence based. More countries were interested in this unique multicultural program. The network founded the EU-Dap Faculty and issues mandates to coordinate the TOT and manage the translation and adaptation of the program materials. This provided an opportunity to assess variations and improvements to the standard program in numerous countries. Figure 14.2 shows where adaptations have started to date.

By the European Commission Public Health Agency as well as through the EU Justice, Freedom and Security and Drug Information projects, EU-Dap had the ability to adapt and implement Unplugged in a large number of EU member states. The United Nations Office on Drugs and Crime (UNODC) regional office for the Middle East and North Africa cofunded dissemination of Unplugged in the region. The nongovernmental organization (NGO) Mentor International Foundation cosponsored an Arabic adaptation and liaised with the IKEA Foundation (IKEA Social Initiative, www.ikea.com) to fund the dissemination process in the Russian Federation, Kyrgyzstan, Croatia, Lithuania, and Romania. It was a significant task indeed for the EU-Dap Faculty to keep track of adaptations and training quality and for the EU-Dap research group to support process and outcome evaluation in the various countries and cultures.

One of the reasons that the Unplugged program was eligible for translation and adaptation is found in the rules of engagement concerning copyright that were agreed to among the EU-Dap network, locating the Unplugged materials clearly in the public domain, with no copyright allowances or royalties implied in the dissemination.

The decision to adopt Unplugged into a nationwide school curriculum or program also was motivated by the argument that it would not only lead to the improved health of youth but also place national or regional authorities among those that are taking science-based and targeted action to reduce drug use. Adopting Unplugged as an effective CSI program also added credibility to the government authority (EU-Dap Consortium, 2008).

Unplugged consists of 12 lessons on intrapersonal or interpersonal skills, knowledge, and attitude. Each lesson relates to several mediating factors, as depicted in Fig. 14.3 (EU-Dap Consortium, 2010). The program encompasses materials for the teacher and the pupils, teacher training, and parent meetings.

Adaptation Is Not Copy-Paste

From a cultural relativist standpoint, best practices must be open to local adaptation in accordance with culture, religion, and other basic principles. The EU-Dap Faculty noted *changes on the surface level: language and idioms, places, product brands, food, music, . . . and the deeper level: risk and protective factors, gender roles, family and community life, legal system, norms and values. . . . It encountered*

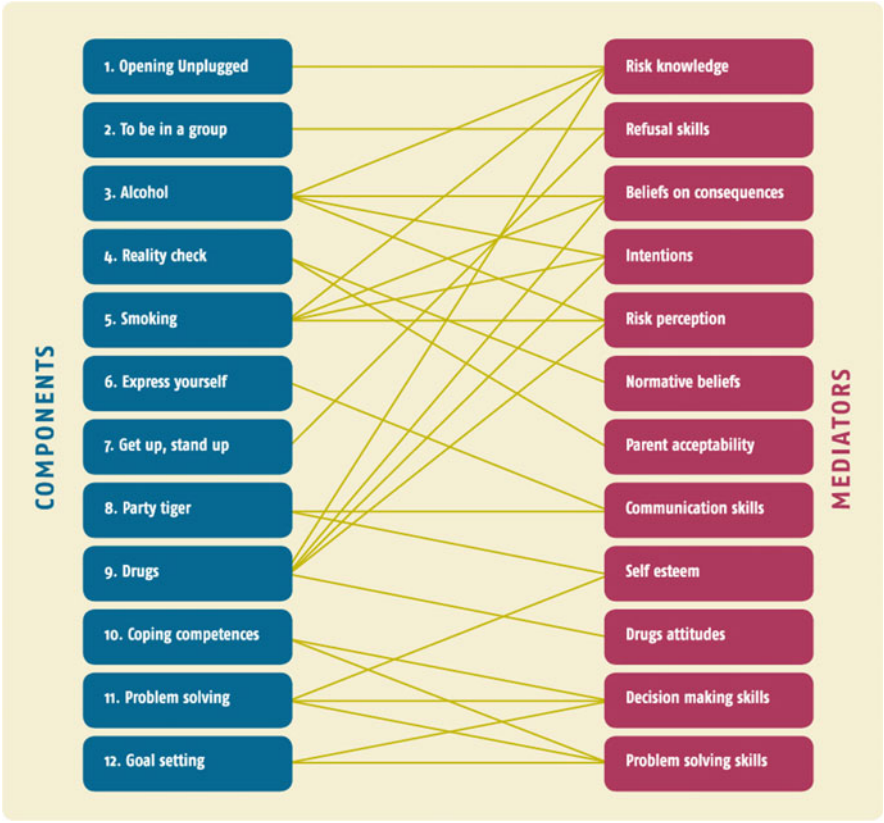

Fig. 14.3 Relation between mediators and 12 Unplugged units

significant changes and adaptations regarding drugs, gender and family (EU-Dap Consortium, 2010). The process compelled EU-Dap to determine the core components of the program, sometimes interpreted in detailed or subtle instructions for exercises. Critical issues identified in Unplugged were social influence and interactive delivery, normative belief or "social norms" related to peer pressure and media influence, and drug-related information in quizzes or information cards. Such issues were part of the program design from the start based on available reviews and meta-analyses (Tobler et al., 2000). *Core components* of the program, however, were determined on stronger fundaments once intermediate factors of effectiveness were analyzed. EU-Dap could classify the components assessed by means of the questionnaire that it had used for the randomized controlled trial in Europe: attitudes and beliefs, social support, school experience, personal skills, knowledge, and intentions (EU-Dap Consortium, 2010). In the adaptation processes, connected to international dissemination, it would become a crucial strategy to distinguish such core components from *variable components*.

The Process of Adapting the EU-Dap Unplugged in the Arab World

The experience of translating and adapting a CSI program with an extensive life skills portion for use in the Arab region contributed to a comparison of very different cultures. The process forced us to outline the distinction between core and variable components more than anywhere else before. As a result of this process, the Arab version of the Unplugged materials have received a final edit and the training module has been delivered in different states. Lessons learned contribute to shaping and improving cultural adaptation processes elsewhere.

Involving Six Countries from the Middle East and North Africa

Over the course of 2009–2011 UNODC Region of the Middle East and North Africa undertook the project "Module Development, Training and Piloting of School-based Education Program for Drug Abuse Prevention" funded by the European Commission program "Promoting good practices and networking for reducing demand for and harm from drugs" – in short: "Unplugged in the Arab region." It overarched the adaptation, training, effectiveness evaluation, and implementation process in five countries: Morocco, Lebanon, Egypt, the United Arab Emirates, and Jordan. Dr. Ghanem Bibi from the Arab Resource Collective (ARC) was commissioned by the NGO Mentor Arabia to translate and adapt the Unplugged materials. ARC has vast experience in "Arabizing" international materials for health promotion. For Unplugged, the basic text Dr. Bibi worked on was a literal translation; he started with a word-by-word check for a good understanding of the English language. Partners from the five countries involved plus a partner in Kuwait had sent feedback on that translation, which was the basis for establishing the adapted version. Feedback was considered per item and country by country; ambivalent issues were debated in a cross-cultural working group with representatives from the Mentor Arabia foundation, EU-Dap Faculty, and UNODC together with the program adapter.

Adaptation Instructions and Conditions

ARC used the instructions set by the EU-Dap Faculty at the kick-off training session and further common sense, judging whether opinions were representative. The given instructions included the following (EU-Dap Consortium, 2010):

- *Start with a literal translation.*
- *Don't report on simple changes like names, sports, places.*
- *There is no need for a full back-translation.*
- *Give an overview of changes you made regarding:*

  - *drug information.*
  - *methodological structure of the lessons.*
  - *more substantial changes in content.*

- *If you want to carry the EU-Dap label, the coordination center, OED, Observatory of Epidemiology on Dependences, University of Turin, Italy, has to be informed of and agree with changes, through the EU-Dap Faculty at the University College Ghent in Belgium.*
- *OED has the copyright on all language versions and locates them in the public domain.*
- *You can also decide to use the program resource independently and go your own way.*

There is more than one Arab language. The Arabic translation ARC started with was standard modern Arabic language because this is generally used in schools. Conditions, however, can still be very different between countries and school classes: local interfaces were needed to "localize" the general version of the materials. Some adaptation discussions focused on using one word instead of another. It was difficult to use one glossary; therefore, the Arab version sometimes includes words from other dialects within brackets. In addition, the actual spoken Arabic language in informal social interactions can be very different from that used in schools or written Arabic. Gaps in educational circumstances (e.g., Internet access for all pupils to 45 pupils sitting in one classroom) also influenced the adaptation efforts.

Checking Adaptation Suggestions to Program Core Components

Regarding content, the ARC tried to keep the original content with respect to user friendliness and community acceptance. In some instructions for role-plays or in situation descriptions, boy-girl relations were replaced by same-sex friendship relations. The EU-Dap Faculty explained, however, that the boy-girl contact in the activities was motivated in part by the intention to have an exchange between a boy and a girl on how to react to peer pressure. Boys can acquire protective strategies from girls and vice versa. This opportunity should not be left out because of cultural differences. One of the role-plays suggested a situation where a boy was doubting whether and how he would approach a girl at a birthday party. The adaptor first suggested replacing the boy-girl interaction with a boy-boy interaction. But taking the above rationale into account, the party-related situation was replaced with a homework-related one, thus allowing the boy-girl communication.

A second repeated change that was proposed by the Arabic adapters was to refer to the illegality of alcohol when this substance was mentioned. The Unplugged materials mention this factor only on a limited number of occasions. The adaptors

proposed mentioning it systematically when alcohol was referred to. However, additional background was supplied for the reason to limit reference to illegality, which is also the case in some European countries where alcohol use is prohibited among those younger than 16 years old. Such information is mentioned in Unplugged only once or twice and is not systematically attached to each alcohol reference. The saturation effect of implicit warnings is one reason, but a more important motivation is that Unplugged focuses on risk factors that are closely related to the lifestyle and immediate concerns of young people: financial cost, aesthetics, and social risks. In consensus the number of references to illegal status in the Arab version was greatly reduced.

Other general comments in the feedback from field workers to the adaptation group included more religious motivations. However, such feedback issues seemed to come from only one NGO or professional in a country to represent a particular vision.

Assignment from Principles to the Particular Detail

Some changes intend to decrease the feeling of insecurity in the teacher. The word for teacher was changed from *lecturer* to a term signifying "someone who forms people, including more sharing." The Arab version adds a chronicle of the European origin and the adaptation process. It explains that the sole Arab version can differ by region. Also, the introduction of the phrase "if this text doesn't apply to your pupils, don't put it aside but try to adapt it to your situation: *think in terms of transfer*" intends to reassure the teacher in his role as deliverer. This instruction is replicated in a delicate Unplugged lesson called "Party Tiger," which deals with self-esteem.

The adaptation group adjoins community, family, and religious organizations as protective factors in the teacher manual. *Father*, *brother*, *teacher* are mentioned more than in the original version as behavior model functions in society. This is again an element of reassurance because it lessens the prominence of illegality as a protective factor. Illegality is still mentioned in the Arab version but is not too pronounced. Illegality, the adaptation group noted, does not have the same significance in the Arab world as in the EU.

Regarding the lesson on drug information titled "Just Say Know," the adaptation group considered mentioning in class at school only those substances that are known to the pupils and that are an issue in society. This is first a decision based on the teacher's assessment and also is dealt with in the teacher training workshops. The information lesson is based on a quiz exercise: pupils receive a few cards with a question on one side and the answer on the other side. A group of two or three students will interrogate another small group. The distribution of quiz cards is an easy way to select the topics for information sharing because cards have different colors corresponding with the substances they are treating (red for alcohol, green for cannabis, etc.). As such the teacher can easily skip cards with substances he

assesses are not an issue. The adaptation group suggested extending in the training course discussion of the cards that are sometimes voluntarily ambiguous precisely to create debate: it seems to be wrong at first sight, but it is right if you think about it longer. Teachers may not like such ambiguity, but the adaptation group emphasized the relevance of moving to this debating style. Translation, editing, and subtle formulation is crucial for that purpose, and the adaptation in this regard led to extra instructions for the TOT.

An issue that occurred many times in the adaptation processes in the Arab countries as well as in the Russian Federation and Eastern European countries is the request to delete references to "possibly positive" effects. Because of a precedent where the government was accused of promoting drugs leading to the death of a young person, the Saint Petersburg adaptation team (EU-Dap Consortium, 2010) was confronted with a formal guideline to avoid any suggestion of possible positive effects related to drugs. The EU-Dap Faculty responded to these requests with the following rationale: referring to "perceived desired effects" is necessary, otherwise young people from target group (12–14 years old) will not understand why other people are consuming substances. Taking into consideration that is it prudent not to be accused of promoting drugs, subtle language is again important: "desired" is a different concept from "positive" although in fact they cover the same effects. Because the comments in Unplugged referring to those desired effects are core components of the program, they cannot be removed. It is necessary to trigger the students think critically and make their own decisions outside of the classroom. A lesson where these effects are discussed helps to correct false impressions pupils may have of the perceived positive effects of drugs. At the same time it instigates an open debate culture in class and adds to the teachers' perceived expertise and trustworthiness on the topic.

Social Context of the Arab Adaptation Process

The institutes involved in adaptation gathered in Beirut, Lebanon, in January 2011, a hinge moment in the region where the "Arabic spring" just commenced. The Arab adapters had thoroughly gone through the program materials, reports, and feedback from trials in classrooms in Morocco, Lebanon, Egypt, Kuwait, Jordan, and the United Arabic Emirates. They assessed that the Unplugged program could be part of a global approach in a fast-changing Arabic society. "How to deal with adolescents" was in this climate a matter that concerned everybody, and in this light Unplugged is perceived as an explosive resource. Social influence, transformation and change, identity, self-esteem are concepts with a highly political connotation in this region. Raising these phrases and concepts within the drug prevention agenda places the program deliverer in conflict with societal or political pressure groups, while political organizations often seem to be happy that prevention professionals are doing their job since it passes the task on to someone else. Moreover, these political organizations do not always discern the effect of an educational or social influence program. Some pressure groups realize what programs of this type are

doing, be it from a defensive perspective: *the people need to be guided*. The adapters also mentioned *enlightened* pressure groups, working from a religious perspective and interpreting the *Koran* to identify how life skills are represented there. The Lebanon adaptation meeting reported observing a movement forward in lifestyle, with alcohol use and knowledge about alcohol as examples from the past decade. Contrary to some reactions of resistance and custody, there was relay of NGOs and educational institutes noticing a momentum for training in critical thinking and life skills. Decision makers and many politicians at that moment, the beginning of the Arabic spring, were eager to act quickly: "*now immediately*."

Adaptation at the Class Level

The fascinating experience of cultural adaptation to the Arab world taught the program designers how content and instructions can change between countries or regions. An adaptation process, however, also takes place at a school level, again with respect to the careful balance between fidelity and change. Cultural sensitivity is not only important on the international or national level of translating, adapting, and printing program materials. Also, within the school a teacher may be dealing with different classroom cultures, requiring adaptation of language, concepts, or approach. What is perfectly fit for debate in one classroom is not necessarily apt to rise as an issue in another. Alcohol drinking as a part of a youth lifestyle, for example, may have to be triggered differently in a class with many Muslim pupils compared to a class with only Western culture pupils. Another class might have experienced a recent death, whereas a second class may comprise only boys. It is the teacher who adapts to the group before him. The teacher must be furnished with the tools and feel secure enough to do so. One instrument is the program materials and instructions, but most applicable is the TOT.

Standardized TOT to Increase Effectiveness

From a methodological perspective, but also with regard to the content of prevention, training is increasingly acknowledged as an essential component for developing schools' and teachers' abilities (Dusenbury et al., 2005; Peters, Kok, Ten Dam, Buijs, & Paulussen, 2009). Training workshops often are received enthusiastically but they face several thresholds. First, there are often no financial incentives attached. Second, teachers are facing increased demands on their workload, while motivation for prevention is lacking when it is not given sufficient support at the political or decision-making level. Last, but not least, teachers often are expected to attend workshops during holidays or weekends. Still, the link of training with specific resource materials may be well received, and when teachers complete training, they are generally motivated and keen to use the program, trying out the freshly acquired methods.

Quality Control in Standardized Training

Quality control in prevention is of increasing importance in Europe, especially as many EU countries are stimulated not only to apply quality programs but also to involve at maximum the local and regional levels in response to drug abuse problems or risks. This note in the European Drug Action Plan 2009–2012 (EMCDDA, 2009) led to a shift of competences for prevention to the local level and delegated responsibility for drug prevention to NGOs. More strategies are pictured to provide common quality criteria or standards at the local level, supporting schools or communities in developing school policies, implementing adequate prevention programs, and assuring minimum quality criteria. Specifications are being defined in some states for the accreditation of prevention agencies or prevention professionals. The standardized training model and certification of trainers and teachers developed by the EU-Dap Faculty are examples of effective quality control instruments (EUDAP Consortium, 2010).

Three Components of the EU-Dap Training

TOT in the EU-Dap Faculty model is designed to be delivered in 2.5 to 3 days (Van der Kreeft et al., 2009). In some countries this is followed up after 6 months to a year with a 1-day booster session for trained teachers. Training modules for teachers include three major components: background and theory, demonstration and experimentation of program activities, and methodology of interactive learning (EU-Dap Faculty, 2011). In the Unplugged training, the three components are given subsequently one after the other but are spread out in small parts over the 3 training days.

Program Background and Theory

The theoretical model and background theory are integrated in the first phase of the training. It can be delivered by lecture or through active group work. An example of the latter is a reading exercise where participating teachers gather in small groups to summarize for each other parts of the background they have read individually. Clarifying questions or comments are captured afterward. The reason to dedicate a considerable part of the training to the theoretical model and theory is that teachers must be able to distinguish core components from variable components of the program. This is necessary for the teacher to integrate and transfer the ready-made program into his existing curriculum, his teaching style, and his school or classroom context. Even when the instruction of the trainer and the intention of the teacher are to implement the program while being loyal to its design, it is inevitable that the teacher

will adapt parts of the program. This requires the teacher to recognize if the core components of the program are neglected in his interpretation. Integrating the theoretical background, philosophy, or vision of the program and a conceptual model in the training also contributes to the necessary study of the program by the initial group of teachers from a school. The training is dedicated to distinguishing core from variable program concepts. It is intended to motivate the participants to further study the program, which is important for the first implementation stage of initiation and occurs again in the last stage of integrating the program into the school context.

Demonstration and Experimentation

A considerable part of the training time is spent preparing, demonstrating, and processing examples of program lessons among the group of teachers. If the group includes 28 participants, a small group of 7 prepares one lesson and presents it to the 21 others, who act as the class group. Such demonstrations are intended for the 21 participants in this case to really experience the program lessons: sit on the pupil's chair. The importance of cautiously building up the risk level, the influence of changing group composition, and the effect of fun and laughter on motivation to cooperate are some examples of didactic elements connected to the life skills method that teachers need to learn by experiment and not by instruction alone (Haystead & Marzano, 2009). A question with a perceived low risk level, such as, "What is the difference between a Facebook friend and a classroom friend?" may seem to be of a relatively higher level than expected once the teacher has to make this distinction with four colleagues in a small group.

Including lesson demonstrations in the training decreases a barrier to implementation: those lessons that have been practiced in training seem to be the lessons that are more often implemented in school afterward. The same counts for standalone exercises: when they have been part of the training or the follow-up, they have more of a chance of being applied in the classroom (EU-Dap Consortium, 2010).

Preparing and processing demonstrations in a group is intended to set the scene for a team approach in schools. Such a team strategy is necessary for successful implementation of any new school program. The training often introduces group preparation of a lesson and sharing experiences on teaching style and technique for the first time since the participants' own education.

Methodology of Interactive Learning

A class where social influence lessons are delivered should not be tranquil. Life skills and social influence lessons include movement, group games, and play. Some skills need to be introduced or developed by using games. Meta-analysis shows that including interactive games in a teacher's method of instruction was effective (Haystead & Marzano, 2009). Teachers are, however, reluctant to apply games. The TOT in the EU-Dap Faculty embraces a lot of interactive game methods.

While exercising energizers and games, workshop participants express experiencing cultural barriers to movement and play. The trainer anticipates this in his model by joining in the game with the group. He expresses his own experiences with pupils or with previously participating teachers. This facilitates for other participants the process of crossing a barrier. The trainer refers to questions and anxieties he notices among the other teachers. Participants tend to reduce their resistance to games and play if they experience such reflections from the trainer. Giving a teacher the feeling that he is not alone contributes to the *readiness to change*, which is imperative for the implementation process.

The interactive component in the TOT is required to guarantee a quality extent of interactivity at classroom-level delivery of a CSI program. Therefore the rationale is explained to the participants: if we intend to increase the positive social influence from one pupil to another at a later age, we must start to exercise this in activities where they are compelled to talk to and work with each other without the teacher intervening. The teacher in this method monitors the process but does not control what is being said. *Interactive* is, in this understanding, more than the teacher going into lively debate with the class and necessarily includes pupil-to-pupil interaction. The impact of this method obviously is bigger when experienced by the teachers during their training than when they are merely instructed to do so.

Implementation and Change

Training and adaptation are factors of standardized and qualitative program implementation. The Unplugged field experiences in and beyond Europe have shown their importance in different cultural settings. The lessons learned also contribute to better implementation at the school level, even at the level of teacher–pupil or pupil–pupil contact. Adopting a new program in a school is always a change and should be regarded in this perspective. The Prevention and Evaluation Resources Kit from the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA, 2010) supports policymakers and prevention coordinators in planning and delivering interventions. It teaches how to shape instructions to school teams and to apply and implement evidence-based programs. Going step by step, starting with a needs assessment and continuing careful study of the program before delivering it to the pupils, seems to be not a trivial undertaking. An introducing example depicts how such logical phases are in reality often completely forgotten.

Frustrating Social Process

Implementation of a prevention program is often considered a *social process*, to be engaged with by a professional team. This is however, not a quarantee to success. At the worst, it may even have adverse effects.

A school principal heard about Unplugged and uses the good social conditions at his school to get started with the program. Three enthusiast teachers with a pioneer spirit and experienced in kicking off new programs are sent to a workshop. The process starts with wild enthusiasm: excellent training, new people, lots of laughs, a great trainer. Back in the classroom, however, the reality is much different because the examples from the training don't work there and the colleagues who attended the team meeting are not interested in applying the concepts in their lessons. It leads to an inert attitude, gratified by frustration. "Who is to blame?" surpasses "What factors disturb implementation?"

The events intentionally were set in motion by the principal, who chose team members with the greatest pioneer spirit for the *social process*. When scrutinized, however, it is a countereffective process: the noble intention of the school principal is not realized at all. Considering implementation as a social process alone does not end well. Quality prevention must be more than selecting a good team.

Program Implementation as Part of the School Change Process

Implementation of a program works better if it is integrated in the context of the school, taking into account other changes occurring or planned. A school is an organization in constant change, and implementing a new program should be inserted into this process. Examining the new program during the initiation period, adapting it during implementation, and cautiously integrating it in the structure are elements of good practice (Van der Kreeft, 2001).

Scrutinize the Program to Get Initiated

When preparing the program for local adoption, a first step should be the close examination of the program materials. In this phase a pilot group is composed, but not one that is based on only the idea that they should form a good team, as in the earlier example. It can be appropriate to include a school counselor, a parent, or even an explicit opponent of such programs because they contribute to the anticipation of future cooperation and debate. This pilot group attends the first training to get acquainted with the program. Note that without neglecting the pivotal function of enthusiasm, this first training should not be considered a sales pitch.

After the training workshop, the pilot teachers need time and space to try out the lessons in their own context. This is imperative to the initiation: tryout lessons allow teachers to become better acquainted with the program and to have an *improved understanding* of it (EMCDDA, 2010). Teaching CSI program lessons might be off-putting to teachers at first. Providing the opportunity to try out the program with no strings attached allows teachers to feel out their own boundaries and those of their class group, building up confidence and actually acquiring the new, often unfamiliar, teaching methods.

Adapt Wisely Without Changing Core Components

Whereas during initiation we experiment with the program units, combining the program and the regular curriculum or dividing the package among several teachers, in the second phase focus should be on how to structurally integrate the program at the school, making adoption of the program mandatory.

To be effective, the evidence-based program must be delivered as it is designed, with all the core elements included. This principle may be confronted with assessments made by program deliverers. From a didactical point of view, good lesson preparation involves assessing the initial situation of the school, the class, the rest of the curriculum, and not least the teacher's own capacities. If the program instructions or the teacher training do not provide a clear determination of core and variable elements, schools tend to identify program core elements themselves. *Adapt wisely* can only be fulfilled if the implementation guidelines contain instructions for the deliverer on how to do so.

Program Sustainability

Once the school has made the CSI program a part of its own puzzle, it can say the program is adopted with respect to the school specificity. The new program should be cemented with a place in the school curriculum. It has to be integrated in the school work plan and the yearly schedule.

The entire school team was informed about what was going on in previous stages, but it is in this stage that the program has to be accepted and acknowledged by the staff as well as by the board or other decisive organizations. The goals of the program, uniquely designed for this particular school, must be well understood by the team. The program content, philosophy, and structure also must be documented and accepted: the CSI program is described on the school website.

Challenges in Implementation: Typical Errors

In addition to the frustrating experience described in the earlier example, some illustrative errors can be anticipated.

- Error 1: Participants of a training try to convince their colleagues of the value of the program soon after the workshop: seperately trying to achieve as home division. But their colleagues did not try out the program in their own environment.
- Error 2: The funders or organizers of the training session require concrete results a few months after the workshop: they expect an immediate return on investment. But the concrete objectives and goals for the program in this particular school can be only delivered when the whole team has become involved.
- Error 3: During the initiation, while the pilot group is becoming acquainted with the program elements, the school brochure or website proudly announces the

introduction of a new prevention program. But at this stage we don't even know if the team will accept its application.
- Error 4: The school applies the program without any adaptation to the specificity of the school, local habits, or regional cultural traits.
- Error 5: A teacher makes his own puzzle by selecting easy-to-apply activities and hence neglects some core program elements.
- Error 6: Adaptations at the regional, school, or teacher level are not documented and not transferred to new teachers, becoming a waste of expertise and energy.

Recommendations

Prevention interventions where the deliverer plays a major role, such as a teacher in a school-based strategy, contain intermediate steps between introducing the program to the program developer and dissemination to the final target group. It is then necessary for the developer or designer to include instruction and description of implementation conditions. For complex interventions like a CSI program, this will include the differentiation of core elements from elements that are open to cultural, local, or even group-level adaptation.

Complex prevention interventions tend to require a TOT or teacher training component. A quality training outline includes theoretical background, practical demonstrations, and interactive delivery. It aims at building the participant's capacity not only to deliver the program activities as they are designed but also to assess the target group composition and situation to make small adaptations. If mediating factors of the intervention are assessed, they are of great value for this capacity at a teacher level. Program trainers should capture and document such instructions in the training materials.

Practitioners and researchers must work hand in hand when developing, implementing, and disseminating evidence-based prevention interventions. The last stage of implementation, where the teacher meets the pupil, or even later, at the stage of pupil–pupil contact, inevitably comprises particular factors that are different from those in the experimental condition. Action-research design therefore best includes implementation conditions distinguishing core factors, mediating factors (if determined), and factors that can or even should be adapted to fit the final target population.

Steps Ahead

The future of the EU-Dap partnership offers fascinating challenges. Further analysis of the database at hand can show with more power which mediating factors are contributing to the effectiveness of the EU-DAP project. A result that is already known is the improvement of refusal skills. New regions following the same study protocol and using the same evaluation instruments can enlarge the sample and

power of such results. Another opportunity to perform a parallel mediation analysis of two comparable programs in the EU and the US is even more promising. It will be interesting to see what is the influence of deliberately not being explicit about these refusal skills in implementation elements such as promotion folders or teacher manuals.

Ongoing adaptation of the Unplugged school-based drug prevention program will lead to improvement of the program content and may contribute components such as booster sessions or extra lessons. Implementation and evaluation in the Southeast Asian Mekong region, or in Brasil, can deliver more cultural adaptations and at the same time give more evidence of core program components. Combining the EU-Dap program with other evidence-based programs such as the Strengthening Families Program (Kumpfer et al., 2008) and studying the effects of such a combination can give direction to international dissemination and coordination of the interventions.

The pools of professional trainers that many evidence-based programs deliver are resources that can be used in a structure of synergy and cooperation. If we join forces and invest in international cooperation and in the expertise and art of prevention trainers, we can contribute to bringing the outcomes of powerful research findings to the capacity-building workshops of the program deliverers and through these field experts teach them to the adult or young target groups of our programs, empowering them to make healthy choices.

References

Berkel, C., Mauricio, A., Schoenfelder, E., & Sandler, I. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science, 12*(1), 23–33.

Castro, F., Barrera, M., Jr., & Holleran Steiker, L. (2011). A critical analysis of approaches to the development of preeventive interventions for subcultural groups. *American Journal of Community Psychology, 48*, 439–454.

Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327–350.

Dusenbury, L., Brannigan, R., et al. (2005). Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research, 20*(3), 308–313.

EMCDDA. (2009). *Annual report on the state of the drugs problem in Europe*. Luxembourg: The Publications Office of the European Union.

EMCDDA. (2010). *Manuals no 4: Prevention and evaluation resources kit (PERK)*. Luxembourg: The Publications Office of the European Union.

EU-Dap Consortium. (2008). *Preventing substance abuse among students, a guide to successful implementation of comprehensive social influence (CSI) curricula in schools*. Turin: OED Osservatorio Epidemiologico Delle Dipendenze.

EU-Dap Consortium. (2010). *The EU-Dap 2 project, final technical report*. Turin: Osservatorio Epidemiologico delle Dipendenze.

EU-Dap Faculty. (2011). *Unplugged teacher's workbook, a guide for the unplugged teacher training*. Ghent: University College Ghent.

Faggiano, F., Vigna-Taglianti, F., Burkhart, G., Bohrn, K., Cuomo, L., Gregori, D., et al. (2010). The effectiveness of a school-based substance abuse prevention program: 18 Month follow-up of the EU-Dap cluster randomized controlled trial. *Drug and Alcohol Dependence, 108*(1–2), 56–64.

Haystead, M. W., & Marzano, R. J. (2009). *Meta-analytic synthesis of studies conducted at Marzano Research Laboratory on Instructional Strategies*. Englewood, CO: Marzano Research Laboratory.

Kelly, J. A., Heckman, T. G., Stevenson, L. Y., Williams, P. N., Ertl, T., Hays, R. B., et al. (2000). Transfer of research-based HIV prevention interventions to community service providers: Fidelity and adaptation. *AIDS Education and Prevention, 12*(Suppl A), 87–98.

Kumpfer, K. L., Pinyuchon, M., de Melo, A. T., et al. (2008). Cultural adaptation process for international dissemination of the strengthening families program. *Evaluation and the Health Professions, 31*(2), 226–239.

Peters, L. W., Kok, G., Ten Dam, G. T., Buijs, G. J., & Paulussen, T. G. (2009). Effective elements of school health promotion across behavioral domains: A systematic review of reviews. *BMC Public Health, 9*, 182.

Tobler, N., Roona, M., Ochshorn, P., Marschall, D., Streke, A., & Stackpole, K. (2000). School-based adolscent drug prevention programs: 1998 meta-analysis. *The Journal of Primary Prevention, 20*(4), 275–336.

Van der Kreeft, P. (2001). *Een gat in de haag: Revision of A hole in the fence and En la huerta con mis amigas*. Merelbeke: De Sleutel.

Van der Kreeft, P., Wiborg, G., Galanti, M. R., Siliquini, R., Bohrn, K., Scatigna, M., et al. (2009). "Unplugged": A new European school program against substance abuse. *Drugs: Education, Prevention and Policy, 16*(2), 167–181.

Chapter 15
Measuring Fidelity

William B. Hansen

Introduction

There is evidence that fidelity moderates program outcomes. With the widespread dissemination of evidence-based programs, a major challenge is to assess the fidelity with which interventions are delivered. Fidelity is a multidimensional construct. Knowing which features of fidelity moderate intervention effectiveness is important to understand how programs achieve or fail to achieve their goals. A key issue, therefore, is the availability of tools for assessing fidelity that can be readily and reliably applied. It will be helpful to advance our understanding of program effectiveness to have methods that can be applied in similar ways across multiple interventions.

Research on fidelity to the implementation of any social program began in the late 1970s with the publication of a RAND report that analyzed federal programs supporting educational innovation (Hall & Loucks, 1977). Since then, researchers have identified multiple dimensions to describe how programs developed in a research context are delivered (Dane & Schneider, 1998). While there is growing research that addresses fidelity in the normal educational environment and in clinical settings, much of the research on fidelity has been developed and applied in school-based settings that mostly focus on substance use prevention. Among the basic concepts important in understanding the quality of delivery are (1) adherence—the degree to which facilitators follow program methods and complete the delivery as outlined in a manual or a curriculum guide, (2) dosage—providing sufficient exposure to the program, often thought of as either the number of sessions delivered or the number of targeted participants reached, (3) quality of delivery—the degree to which the person delivering the intervention executes with skill and understanding, (4) engagement—the degree to which

W.B. Hansen (✉)
Tanglewood Research, Inc., Greensboro, NC, USA
e-mail: billhansen@tanglewood.net

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_15,

participants are appropriately involved in intervention tasks, and (5) program differentiation—the degree to which the intervention achieves immediate and intermediate outcomes that differentiate implementation from either a situation in which there is no intervention or an alternative intervention that targets different immediate and intermediate outcomes (Dane & Schneider, 1998; Dusenbury, Brannigan, Falco, & Hansen, 2003).

Durlak and DuPre (2008) note an additional dimension of fidelity that should be considered. Adaptation—the degree to which a program's methods and content are modified, including developmental and cultural appropriateness—is related to fidelity but is usually considered conceptually different from fidelity (Dusenbury et al., 2003). Our discussion includes the measurement of adaptations. Durlak and DuPre also note that, at least in a research context, monitoring of control or comparison conditions to assess any treatment contamination or background intervention service delivery should be considered. Finally, program reach that includes participation rates and the degree to which the intervention is delivered to intended target audiences is worthy of attention. Because monitoring of nonintervention groups and program reach have not been studied, a discussion of these is not included.

In this section, the methods used to assess each of the six dimensions of fidelity mentioned above are reviewed and critiqued. It should be noted that the variety of methods presented in the literature to date suggests that methods for assessing fidelity have not yet been fully developed, tested, or standardized (Mowbrey, Holter, Teague, & Bybee, 2003). Indeed, O'Donnell (2007) concludes that too few methods have been developed, and among studies where fidelity has been measured, goals and methods have varied considerably.

An issue that has not been systematically researched, but is undoubtedly one that requires attention when collecting fidelity data, is the challenge of terminology. When observers and coders are trained and as they perform their duties, the operational definition of terms comes to predominate as something that needs to be mastered (McGrew, Bond, Dietzen, & Salyers, 1994). The language of fidelity assessment is in no way universally applied nor is it universally understood. Quite the contrary, many of the terms researchers, program developers, evaluators, and policy makers use end up having quite specific and sometimes idiosyncratic meanings. Even the most basic of terms—*fidelity*, *adherence*, *dosage*, *engagement*, *program differentiation*, and *adaptation*—may have one meaning for an evaluation staff and a very different meaning for practitioners. This places the development of suitable self-report measures on the horns of a dilemma. On the one hand, the language needs to be sufficiently specific to allow responses to fit within a pre-specified schema. On the other, the language cannot be so cumbersome in an effort to be precise that practitioners give up trying to understand it and then fail to provide meaningful data. For instance, few, if any, teachers or clinicians have the time to read a coding manual. If self-reports are asked for, they need to assess information these individuals can realistically provide and use language that can be easily and unambiguously understood in a format that they can follow.

It should be noted at the outset that there are two approaches to how fidelity assessments are structured. One approach to measurement examines highly

structured programs such as multisession school-based interventions. These usually involve assessing program implementation against manual-based instructions that direct in some step-by-step manner how an intervention is prescribed to occur. However, when measuring fidelity in clinical settings and with environmental interventions, rather than structured multisession interventions, other factors are the focus of measurement. These two settings—namely, with manual-based instructions and without manual instructions (clinical settings and environmental interventions)—each require a different approach to measurement. Our examination of fidelity measurement will address each approach separately.

Measuring the Fidelity of Manual-Based Interventions

This section discusses measuring the fidelity of delivery of interventions that have some form of prescribed step-by-step procedures. One of the characteristics of the interventions is that they tend to be divided into discrete sessions, each of which has some level of detail describing specific goals or objectives to be achieved and procedures specific to the session that should be followed. That is, in general, the sequencing of the sessions is prescribed, all to be delivered within some defined period.

Content of Fidelity Assessment: Measurements

The grist of fidelity assessment, especially when applied to manual-based programs, has to do with what is actually measured. Each of the six broad subcategories of fidelity has been addressed by research using a variety of methodologies for collecting data. However, across all methodologies, common elements emerge that are summarized here.

Adherence. Adherence is the most basic component of fidelity and has been measured in several ways when structured multisession programs are studied. When observers provide data about adherence, they typically document (1) the percentage of points and objectives covered (Botvin, Baker, Dusenbury, Tortu, & Botvin, 1990; Botvin, Dusenbury, Baker, James-Ortiz, & Kerner, 1989), (2) coverage of key program concepts (Spoth, Redmond, & Shin, 1998), or (3) the extent to which the teacher followed the protocol (Botvin, Baker, Filazzola, & Botvin, 1990; Sobol et al., 1989).

For example, for each session of the cadre of *LifeSkills Training* programs, there is a downloadable fidelity checklist (http://www.lifeskillstraining.com/lst_process_tools.php) that consists of "yes" and "no" response options to whether "major objective(s) and corresponding points" were completed as well as "yes" and "no" response options to whether "topics and activities" were covered. Spoth and colleagues (Spoth, Guyll, Lillehoj, Redmond, & Greenberg, 2007, 2011, Spoth,

Guyll, Redmond, Greenberg, & Feinberg, 2011) used a similar method for a parenting program and three school-based programs (*LifeSkills Training*, *Project ALERT*, and *All Stars*). Observers completed adherence checklists, marking whether facilitators delivered specific content and activities prescribed by the intervention. In this study, the adherence score was calculated as the proportion of content delivered in each session.

The primary difference in how adherence measures have been collected by observers and teachers is on the level of detail requested from each group. Both groups can provide information about fulfillment of goals or objectives as well as about completion of sessions and activities. When teachers have provided these measures, they have described the (1) percentage of content delivered including percentage of sessions implemented (Basen-Enquist et al., 1994; Pentz et al., 1990), (2) percentage of points covered (Dent et al., 1998), and (3) teaching strategies used (Tappe, Galer-Unti, & Bailey 1995).

There have been studies in which the same adherence measures were collected by both observers and intervention facilitators. For example, Breitenstein and colleagues (2010) assessed adherence to the delivery of a parent training intervention delivered to low-income families of small children, the *Chicago Parent Program*. To create the adherence scale, they performed a content analysis of the intervention and, for each of the 12 sessions, created yes/no checklists of between 12 and 16 essential elements of the intervention prescribed for delivery. Observers used audio recordings to make judgments. Facilitators provided post-session self-reports. Paired observers had high rates of agreement about which elements were delivered (94 % agreement). Observers overall found that 89 % of essential elements of the intervention were delivered. Average agreement between observer and facilitator adherence ratings was 85 %. Interestingly, most of the disagreements (87 %) occurred when facilitators claimed to have delivered content and observers noted that the content had been skipped. As noted earlier, this self-report bias is expected.

Ennett and colleagues (2012) differentiated between two measures of adherence: content delivered and using prescribed delivery strategies. Content questions were tailored to the logic model of the intervention and, depending on the curriculum being assessed, included assessments of the degree to which teachers addressed information (such as drug use consequences), refusal skills, personal competency skills, positive affect, beliefs, and attitudes.

Dosage. Projects that have measured dosage have generally adopted some variant of a method for documenting the overall exposure to the program. Thus, researchers have measured (1) the total number of subject areas taught (Basen-Engquist et al., 1994), (2) students' exposure across the complete intervention based on records of attendance (Dent et al., 1998), (3) time spent completing the program (Pentz et al., 1990), (4) sessions implemented (Ennett et al., 2012; Rohrbach, Graham, & Hansen, 1993), and (5) completeness of implementation (Tappe et al., 1994). Each measure provides a sense of how much programming was delivered.

For example, in assessing *PATHS* (Greenberg, Kusche, Cook, & Quamma, 1995; Ransford et al., 2009), teachers were asked, "How often, on average, do you actually use the PATHS curriculum lessons and generalization techniques in your classroom?" Possible responses included *not at all*, *rarely*, *only when problems arise, occasionally, every few months, regularly, 1–2 lessons per week, frequently,* and in *weekly lessons with frequent generalization techniques*. This measure mixes dosage and generalizing and, as such, lacks the parsimony typically expected of measures of dosage. Nonetheless, average number of lessons taught and average number of supplemental activities taught were significantly related to receiving coaching, $\beta = 0.26$ and $\beta = 0.32$, respectively. Although there are no data to assess the reliability of self-reports, both measures reflect qualities that might reasonably be expected to increase as a result of increased encouragement and training.

From the perspective of program facilitators, dosage often translates into a measure of how many sessions of an intervention were delivered or how much time was devoted to delivery. Observers are often limited in their ability to assess this notion of dosage. Although there are exceptions, this limited view is often imposed because observers only sample sessions and rarely view the intervention in its entirety. This, of course, can be dealt with as has been reported by Sloboda and colleagues (2009) by knowing when sessions are scheduled and scheduling observers to be present for all sessions. However, observers are much more likely to focus on such elements of dosage as class size (Fagan & Mihalic, 2003) and attendance (Gottfredson et al., 2009).

Quality of Delivery. Quality of delivery is a more subjective aspect of fidelity to measure. Both observational and self-report methods for measuring quality of delivery can be informative. In a survey of special and general education journals, Swanson, Wanzek, Haring, Ciullo, and McCulley (2011) found that only 9.8 % of the studies they reviewed included information about the quality of intervention delivery. However, it might be noted that this topic is among the most widely divergent in terms of the measures employed. The types of measures researcher use have included (1) assessing the general quality of teaching practices (Abbott et al., 1998), (2) ratings of teacher effectiveness and enthusiasm (Botvin et al., 1989; Hansen, Graham, Wolkenstein, & Rohrbach, 1991), (3) assessments of the quality of the teaching strategy used (Giles et al., 2008; Harachi et al., 1999), and (4) ratings of the instructor's performance in being able to involve participants in discussion (Sobol et al., 1989).

For example, Giles and colleagues (2008) assessed quality of delivery of *All Stars* by asking observers to note teachers' propensity to praise and encourage students, accept students' ideas, incorporate students' ideas in teaching, ask original questions not asked for in the curriculum, repeat questions to encourage multiple students to respond, ask probing questions, and manage their classroom. Teachers provided video recordings of class sessions. As observers watched these videos, they counted occurrences of each quality. Each measure resulted in an intra-class correlation between paired observers that was greater than 0.85, except in the case of incorporating students' ideas (ICC = 0.73). The challenge with this method is that it must be standardized to the curriculum, which means that counts may vary

from curriculum to curriculum based on how each is structured. As yet, cross-intervention comparisons have not been completed, and this method awaits further testing.

A similar method was used by Breitenstein and colleagues (2010). Observers listened to audio recordings of program sessions and rated facilitators' competence at delivery using 15 criteria, each of which was rated on a 3-point scale that assessed teaching skills. An example includes the following: "The facilitator correctly conveys and communicates program principles." Paired observers had on average 85 % agreement on these ratings. Possible responses included *skill rarely or never demonstrated*, *skills emerging, skill that needs further development*, and *skill demonstrated and done well*. Whereas Breitenstein's group had asked facilitators to assess their own adherence to the program, as was described earlier, there was no attempt to ask them to rate their own competence.

Ransford and colleagues (2009) assessed quality of delivery through teacher self-reports by asking teachers, "How well do you feel you are implementing the lessons in the PATHS manual?" Ratings included *not at all*, *not very well*, *somewhat*, *fairly well*, and *very well*. Teachers had higher quality of delivery ratings if they felt that they had administrative support and coaching during implementation, both of which might suggest that this measure is a sensitive indicator of quality.

Another way to approach quality of delivery is to ask teachers to estimate their effectiveness or efficacy. For example, Ransford and colleagues (2009) asked teachers to complete a 15-item questionnaire about their efficacy that asked teachers to estimate how effective they were as teachers. Questions included items such as "If students stop working in class, I can usually find a way to get them back on track." Response categories ranged from *strongly disagree* to *strongly agree* using a 6-point scale with *moderately* and *slightly* filling in the other ratings. Inter-item reliability was 0.64 for the combined items.

One aspect of fidelity that has not apparently been the focus of measurement but certainly fits within the concept of quality of delivery is the degree to which facilitators make identifiable mistakes in intervention delivery. That is, there are as yet no methods for assessing when a teacher presents a contradictory or a contraindicated message. For example, it is not uncommon for teachers who deliver interventions that focus on establishing low-use norms about drug use to send the messages that drug use is quite prevalent (Dusenbury et al., 2005).

Engagement. Tobler and Stratton (1997) identified interactivity to be important in achieving prevention outcomes particularly for middle school-aged children. Generally speaking, engagement is important because many of the mediating effects targeted by interventions focus on attitudes, intentions, beliefs, and skills that typically require participants to internalize concepts. That is, one of the essential features of a successful intervention is that it promotes such internalized changes. Tobler and Stratton characterize engagement and interactivity as a structural component of an intervention (e.g., the calling for role-plays, games, or small group activities), which not all interventions shared. Giles and colleagues (2008) instead focus on the degree to which interactivity and engagement appear when any intervention is delivered. Thus, when used in the context of assessing fidelity,

interactivity (or participant engagement) is a dynamic component—the degree to which teachers actively engage students as called for by the program they deliver. Engagement has been measured in a number of ways using both observers and teachers as sources of data. The types of measures researchers use have included an overall subjective rating of quality of delivery and participant engagement (Hansen, 1996; Pentz et al., 1990).

Specific examples of strategies to assess participant engagement include those used by Spoth and colleagues (2007) in a parenting program. Their method called for observers to give 0–4 ratings for the degree to which participants were engaged and the amount of active family participation that was observed. This was assessed with four ratings that included (1) students' attitudes toward the lesson, (2) the appropriateness of their behavior, (3) their willingness to engage in discussion, and (4) their manifest interest in program material and activities.

Giles, Harrington, and Fearnow-Kenney (2001) used student reports to assess their engagement in the delivery of *All Stars*. This measure included six questions asked at posttest and included such prompts as "I paid attention during *All Stars*," "I liked the one-on-ones" (meetings held between teacher and student), and "I liked the questions my teachers asked in *All Stars*." This scale had high internal consistency (alpha = 0.93) as measured in a later study (Ringwalt et al., 2009). In a subsequent research study that used observational methods applied to video-recorded observations (Giles et al., 2008), coders counted occurrences of seven categories of teachers' behaviors. These included (1) praising and encouraging students, (2) accepting and using students' ideas, (3) asking questions, (4) disclosing personal experiences and examples, (5) managing the classroom, (6) lecturing (negatively coded), and (7) giving directions. All measures had intra-class correlations between paired raters greater than 0.83.

In more recent research (Bishop et al., 2013), paired observers rated video recordings of teachers delivering *All Stars* using a single measure of engagement. Coders were instructed to assess how directly students participated in session activities as well as the proportion of students involved in activities. This measure resulted in only modest inter-rater reliability (ICC = 0.60) and was not deemed sufficiently reliable to recommend it in other research. Instead, the authors suggest assessing student engagement directly from students using posttest measures similar to those examined by Giles and colleagues (2001) or using multicomponent analyses (Giles et al., 2008; Spoth et al., 2007).

Program Differentiation. This area of fidelity has been the hardest for researchers to define, as evident by the relative lack of studies that have attempted to assess this dimension. In part, the challenge lies with interpreting the original definition of Dane and Schneider (1998). Their original conceptualization—unique features of the intervention are distinguishable from other programs—has had several interpretations. For example, Ennett and colleagues' (2012) assessment of a broad sampling of school-based program implementations ascribed the use of just one evidence-based intervention as providing evidence of program differentiation. When two or more programs were implemented, program differentiation was said to have disappeared.

A separate approach to assessing program differentiation has relied on mediating variable analysis, an approach that calculates the degree to which interventions alter mediators which, in turn, account for changes in targeted outcomes (MacKinnon, Fairchild, & Fritz, 2007). A study that examined this issue (Hansen et al., 1988) compared the impact of three different alcohol curricula—information, resistance skill training, and normative education—on targeted mediating variables. Students in all conditions were assessed on the same battery of mediators, which included their knowledge about the effects of alcohol, their belief in their self-efficacy to resist peer pressure, and their beliefs about the prevalence and acceptability of alcohol use among peers. Results showed that each intervention produced desired results that were unique to the program and did not overlap with changes observed in each of the other programs. That is, the informational intervention uniquely increased students' knowledge about the effects of drinking alcohol but had no effect on either their ability to resist peer pressure or their understanding of peer norms. Similarly, the resistance skill training intervention increased students' skills for resisting peer pressure but had no effect on knowledge about consequences and did not improve normative beliefs. Finally, the normative education intervention altered students' beliefs about the prevalence and social acceptability of alcohol use but did not alter knowledge or self-efficacy.

This same basic idea about the role played by mediators can be translated to all interventions. That is, once a logic model is specified for an intervention that includes either defined immediate or intermediate effects, the impact of the intervention on these then serves as the basis for assessing program differentiation. In these cases, there must be a counterfactual—either a control/comparison group or an alternative intervention that has distinctly different expected immediate or intermediate outcomes. An alternative is to compare multiple cases (e.g., schools or classrooms) to distinguish those for which positive outcomes are observed with those for which intended outcomes are not observed.

In school- and community-based settings, these outcomes are most often measured by including mediating variable indicators in surveys, completed at pretest and posttest. In addition to the example provided above, other examples include research by Baranowski, Anderson, and Carmack (1998); Baranowski, Lin, Wetter, Resnicow, and Hearn (1997); Botvin, Baker, Dusenbury, Botvin, and Diaz (1995); Donaldson (2001); Hansen and McNeal (1997); MacKinnon and colleagues (1991); McNeal, Hansen, Harrington, and Giles (2004); and Teasdale, Stephens, Sloboda, Grey, and Stephens (2009).

Adaptation. Adaptation is the least studied fidelity construct. Many prevention researchers have conceptualized adherence and adaptation to be opposite ends of a single continuum. Indeed, of the projects reviewed by Dusenbury and colleagues (2003), those that measured adaptation did so primarily by noting a lack of adherence. This approach fits with the view of early researchers (e.g., Boruch & Gomez, 1977) who viewed changing any elements of an intervention as essentially diluting its potential for effectiveness. More recently, however, researchers have advocated for a position that allows for reinvention and adaptation short of the "zone of drastic mutation" (e.g., Blakely et al., 1987). Recently, Hill, Maucione,

and Hood (2007) examined adaptations and deletions in a disseminated community-based prevention program. They suggest that assessment should focus on the few types of adaptations that are most widespread. In their research, they identified that a major reason for adaptation and deletion was lack of time during implementation.

The latest evidence suggests that adherence and adaptation may not be an either/or phenomenon. In a study examining *LifeSkills Training* as disseminated (Dusenbury et al., 2005), observers rated adherence in terms of percentage of objectives reached and number of activities completed following Botvin's model (Botvin, Baker, Dusenbury, et al., 1990; Botvin, Baker, Filazzola, & Botvin 1990). Observers also noted the additions and modifications made by teachers as sessions were taught. This content was then analyzed for its potential to enhance or detract from the goals of the program. Results suggest that increased adherence was correlated with making positive adaptations; teachers who were most adherent made positive adaptations, whereas those who were least adherent made adaptations that detracted from the program's goals. This suggests that adaptation should be conceptualized and measured independently of adherence.

A recent study by Hansen and colleagues (2013) examined 306 video recordings of teachers delivering *All Stars*. Observers noted the number and type of adaptation each teacher made. Adaptations that were deleterious to program goals received negative valence ratings and positive ratings were given for adaptations that were likely to facilitate achievement of program goals. Those who made few adaptations, and whose adaptations were rated as having the potential to improve program outcomes, had a higher percentage of students who remained non-drug users compared to teachers who made many or negative adaptations. Students whose teachers either made relatively many adaptations or whose adaptations were negatively rated had significantly greater drug use onset. Measures of fidelity, including quality of delivery and teacher understanding, were related to valence of adaptations, with better performance related to making positive adaptations.

Multiple Measures

Each of these concepts has been examined in empirical studies, although early studies tended to focus on one or two of these dimensions and have not typically addressed all dimensions simultaneously. Ruiz-Primo (2006) proposed a multimodal, multidimensional approach in which each of the four markers of fidelity—adherence, dosage, quality of delivery, and participant responsiveness—was assessed from the perspective of program facilitators, observers, and program participants. Measures were tailored to the ability of the target source of data to provide information. Thus, for example, teachers provided a teaching log to provide information about adherence, observers watched video recordings to gather this information, and students' notebooks were examined and scored to reveal whether the intervention had been delivered.

Fagan and colleagues (2008) collected data as part of the evaluation of 16 evidence-based programs involved in the *Community Youth Development Study* that included measures of adherence, dosage (number, length, and frequency of sessions), participant responsiveness (similar to engagement, with a focus on participation and perceived understanding), quality of delivery (teacher skill), and documentation of implementation challenges. However, rather than using one standardized instrument, the research team adopted methods for assessing fidelity prescribed by each program.

There is value to each approach, and the employment of a combination of measures (e.g., assessing objectives as well as completion) is expected to enhance measurement quality. Hansen and colleagues (Gottfredson et al., 2009; Hansen et al., 2011) have recently developed a system, *Evaluation Lizard*, http://www.evaluationlizard.com that integrates typical adherence, dosage, quality of delivery, engagement, and adaptation into one format for use with the sessions and activities of all session-based drug prevention programs that appear on the Substance Abuse and Mental Health Services Administration's National Registry of Evidence-based Programs and Practices and that can be adapted for the use with other session-based interventions.

Data Collection Methods

The fidelity of manual-based interventions is generally assessed using one or more of five methodologies: (1) using self-reports from the facilitator, teacher, clinician, or intervention service provider; (2) observing and coding video or audio recordings; (3) directly observing (e.g., trained raters visiting the site of intervention); (4) using interviews and surveys of program participants to collect data; and (5) searching archived records. The latter two are rarely the sole source of information when examining manual-based interventions but often supplement data collected using the first three methods. Indeed, the last method, searching archived records, is rarely used for assessing the fidelity of a manual-based intervention.

The quality and type of information gathered using each method may vary considerably, and methods are usually selected based on the objectives to be fulfilled. Self-reports are relatively easy to collect in terms of requirements of time and effort (Ransford et al., 2009). There are some clear benefits to gathering self-reports. Among these are that those who implement programs may be in the best position to report on their attitudes and motivations and provide other insights that cannot be gained from observation. However, the challenge with self-reports is that there is also significant potential for a self-report bias (Adams, Soumerai, Lomas, & Ross-Degnan, 1999; Lillehoj, Griffin, & Spoth, 2004) that may threaten the reliability and validity of any findings. For example, in recent research in which observers and teachers rated their performance using parallel forms, teachers made more positive fidelity ratings than did observers (Hansen, Pankratz, & Bishop, under review). With the exception of ratings about adherence and attendance,

teachers and observers failed to agree on fidelity and adaptation ratings. Teachers' ratings were not predictive of students' scores, whereas observers' ratings were.

Such biases may not always occur, but because they occur relatively frequently, there is a general tendency to use self-report only when other methods cannot be used. Generally, when there is a bias, it tends to favor the teacher and put the teacher in the best possible light. In part this is because self-reports have the potential to be used as the basis for evaluating a teacher's performance (Donaldson & Grant-Vallone, 2002). Further, objectivity about self is not a universal quality. Teachers may not understand the concepts they are reporting about or may not be able to judge their own performance according to an evaluator's standards. The solution to this problem may be to increase the specificity of what is called for with teacher measures or to provide training and examples. However, such measures may only be able to achieve modest gains in self-report qualities. Teachers are pressed with other demands and are likely to see little value in expending the effort required to become expert at completing fidelity assessment forms. Thus, the value of teachers' self-reports may continue to be limited.

On the other hand, when used to supplement other methods of data collection and when accompanied by guarantees of confidentiality, self-report has significant potential to provide useful and meaningful data particularly regarding the instructors' understanding and perception of the intervention as prevention and its appropriateness within the context in which the intervention is delivered. Because observations may need to sample sessions rather than be applied to the entire course of intervention, teachers may be in the best place to report about some things of interest. If multiple session interventions are being evaluated, teachers may be in a better position than observers to provide information about qualities of the intervention as a whole. Teachers and other intervention providers may also be in the best position to provide a commentary about the quality of the intervention itself. They may have a unique understanding of how well directions are given in protocols and manuals and can provide insight into how these directions influence their and their students' performance. They may have unique understanding of the barriers faced by their participants. Indeed, a multi-method approach may yield the highest benefit in terms of gaining a full understanding of implementation fidelity (Ruiz-Primo, 2006).

Video and audio recordings have numerous benefits for the collection of fidelity data. Segments can be paused and replayed if needed. This allows for validity and reliability to be established, with multiple observers able to view exactly the same series of events from the same perspective. This method requires equipment such as camcorders with tripods. The current technology for collecting video recordings with high-quality sound is relatively inexpensive. Further, it is relatively easy to standardize procedures for setting up a recording system so that recordings are made in a routine manner. Digital data files containing images and sound are easy to transmit and to store. There are nonetheless several challenges with using video and audio recordings as the source of data. For example, cameras, particularly when placed in a stationary location, have limited scope and miss viewing or hearing the entire intervention. With video recording, panning and zooming are possible

methods that can be used to capture the dynamics of a setting; however, such techniques are often poorly and inappropriately applied and are not recommended. Additional advantages that recordings afford include the ability to train coders to specific coding criteria and, should coding standards change, to make corrections to prior work.

On the downside, once the session is recorded, that is all there is to go by. If equipment fails or if recordings are not started on time, data are lost. Recordings may also focus attention on minutiae, so should a meta view be desired, coding can become a challenge. For example, recordings are ideal for counting discrete events, but elements such as overall atmosphere and other qualities that require summary judgments may be harder to assess. Finally, there is the possibility that the presence and operation of recording equipment may distract program participants. With care and experience these challenges can be overcome. When making recordings, there are issues related to the protection of teachers' and students' identities that must be considered. Such issues as informed consent by parents, assent from students, and placement of equipment become issues that must be addressed.

Onsite observation is an alternative to coding recorded images and sounds. The value of onsite observation is that it provides a means for assessing the context in which the intervention is delivered, which can be a feature useful for evaluation. Often, onsite observation is done only after extensive training. Observers can be trained to track specific features of an implementation and can adapt their focus to make records or ratings of these. There are, however, several challenges that in situ observers face that may make onsite methods less than ideal for some needs. There is no possibility of pause and replay. If actions within the intervention happen quickly, if words are only somewhat audible, if there are distractions, or if there are issues that require thought, judgment, or discussion, observers may be challenged to make sufficiently clear notations for use in subsequent coding and analysis. Errors are difficult to correct, and there is limited possibility to update observations if coding standards change. As with recorded observations, having an onsite observer may create a distraction. Further, perhaps more than is the case when recording equipment is used, there may be a reactance on the part of the intervention leader who may behave differently under the scrutiny of direct observation. It is often the case that findings from onsite observations are also used to provide feedback to the instructor (McNeal & Hansen, 1999). This can be beneficial if the goal of the observation is to improve performance. However, if the goal of observation is simply data collection, this may be an unwanted by-product. Indeed, it is often the case that, even when the goal is primarily research oriented, teachers will want feedback about their performance, even when not explicitly requested. In the Hansen and McNeal study, observers were instructed to share their findings with teachers. To avoid creating offense when findings were shared, observers were instructed to use only the upper values on 7-point rating scales, essentially reducing the used values to 4–7. In contrast, Sloboda and colleagues (2009) instructed observers to tell instructors that they could not give ratings because final scores needed to be calculated. Instead, observers learned to say something like "This lesson went well" or "I enjoyed this lesson." The key is to understand that feedback

is almost always expected and some strategy for responding to requests for feedback should be planned.

The final common strategy for collecting information about the fidelity of manual-based interventions is to ask program participants themselves about the intervention. Just because interventions are delivered does not always mean that they are received or that they are received as intended. Measures that assess how participants respond to features of an intervention can reveal a great deal about fidelity. In the distant past, participant satisfaction assessments were often the only markers of intervention success that were gathered. Beyond satisfaction, features of fidelity such as engagement and interest, feeling respected, and so forth are worthy of attention. Moreover, to the extent that affecting targeted immediate and intermediate outcomes in participants is considered an essential element of assessing fidelity, participant surveys become a key method to consider.

The specific nature of which qualities related to fidelity are assessed should not be confused with outcome evaluations in which effectiveness is measured in terms of changes in behavior. However, it is nearly universally the case that targeted mediators (modifiable risk and protective factors) can constitute one level of fidelity assessment. These are the classic manipulation checks that tell whether the intended immediate or intermediate changes in attitudes, knowledge, beliefs, values, skills, social relations, and other targeted qualities have benefitted from the intervention. When multiple programs are being delivered, these manipulation checks can be treated as measures of program differentiation. That is, the effects of each program should be specific to that program and there should be no crossover between or among programs. When a single program is being evaluated, such manipulation checks still fall under the general rubric of program differentiation and mediating variable analysis may need to suffice as the measure of differentiation.

Thirty-seven studies that employed these various methods for collecting fidelity data were identified as part of this review (Table 15.1). Although not an exhaustive list, Table 15.1 gives a snapshot of the methods employed by researchers in assessing fidelity. When measuring adherence, observational methods predominate as the source of data. Nonetheless, facilitator self-reports are common and often stand alone as the source of information about adherence. Participant reports were used only once. On the other hand, facilitator self-reports are the most common method for assessing dosage. As will be noted subsequently, dosage when reported by facilitators differs in content from dosage when reported by observers. Examining archival records and gaining information about dosage from participant reports have been used relatively rarely. As was the case with adherence, measuring quality of delivery was most frequently performed by observers. However, facilitators were also frequently asked to provide information about the quality with which they delivered interventions. Participants only rarely provided information about this. In contrast, both observers and participants frequently provided information about how engaging the intervention had been. Program differentiation was primarily the domain of surveys conducted among program participants. How program differentiation was assessed from participants differed markedly from

Table 15.1 Frequency with which various measures of fidelity and their source are reported

		Adherence			Dosage				Quality of delivery			Participant responsiveness			Program differentiation		Adaptation	
First author	Year	F	O	P	F	O	R	P	F	O	P	F	O	P	F	P	F	O
Beets	2008	✓			✓				✓									
Bellg	2004					✓	✓											
Biggs	2008				✓				✓									
Botvin	2006		✓															
Breitenstein	2010	✓	✓							✓								
Domitrovich	2000				✓					✓								
Ennett	2012	✓			✓				✓			✓			✓		✓	
Fagan	2003		✓		✓	✓				✓			✓				✓	✓
Fagan	2011	✓			✓							✓						
Feinberg	2010															✓		
Giles	2001										✓			✓				
Giles	2008								✓					✓		✓		
Gottfredson	2009		✓			✓				✓			✓	✓		✓		
Griffin	2006		✓							✓								
Hanley	2009	✓			✓													
Hansen	1988															✓		
Hansen	1991								✓	✓								
Hansen	2011		✓			✓				✓			✓					✓
Hill	2007																✓	
Johnson	2010		✓			✓				✓								
Lee	2008	✓							✓									
Lillehoj	2004	✓			✓													
MacKinnon	2007															✓		
McGrew	1994					✓	✓											
Mihalic	2008				✓				✓									
Pankratz	2006		✓															✓

Ransford	2009	✓			✓				✓										
Ringwalt	2003																		
Ringwalt	2009										✓			✓			✓		
Rohrbach	1993				✓														
Rohrbach	2007		✓							✓			✓						
Rohrbach	2010									✓			✓	✓					
Ruiz-Primo	2006	✓	✓	✓		✓		✓	✓	✓	✓		✓	✓					
Skara	2005		✓							✓			✓						
Spoth	2007		✓										✓						
Tiwari	2010		✓																
Wilson	2009		✓			✓													
Times observed		9	14	1	11	8	2	1	9	12	3	2	8	6	1	6	3	3	

F Facilitator, *O* Observer, *R* Records, *P* Participant

how it was assessed from intervention providers. Finally, observers and facilitators were equally likely to be asked about adaptations they made to the intervention.

Measuring the Fidelity of Clinical and Environmental Interventions

This section discusses measuring the fidelity of interventions that are either clinical or environmental in nature. The primary characteristic that sets these apart from manual-based interventions is that, instead of having discrete parts, the intent is to provide an intervention that is continuous and persistent. So, instead of asking whether a step has been completed related to achieving an objective, the goal for measuring fidelity is to understand the degree to which a system is in place and has ongoing functions (Bond et al., 2000).

Data Collection Methods

Methods for collecting information about the fidelity of clinical and environmental interventions rely primarily on archival and observational sources, with much less attention paid to collecting surveys and self-reports. Thus, hospital records, records from the criminal justice system, employment records, and other archived data are mined for information that might provide evidence of what the nature of the intervention has been. Site visits, policy document reviews, and financial reports can all provide useful information. In a crucial sense, the goal of collecting data is to provide a basic description of the system that is in place and any data that can be used to document the state of affairs regarding the intervention becomes useful.

The Content of Fidelity Assessment

The term *fidelity* takes on a different meaning for clinical and environmental interventions. For clinical and environmental interventions, fidelity refers to the degree to which the system approximates a prescribed model. Therefore, fidelity in this setting focuses on the degree to which the system replicates the treatment model. Before assessing fidelity, it is crucial to have an understanding of the treatment system or the environmental features that are expected. In these senses, fidelity becomes interchangeable with adherence.

An example of how this is accomplished is described by McGrew and colleagues (1994; see also Teague, Bond, & Drake, 1998) who measured the fidelity with which community mental health services were provided. The first task set about by

the research team was to get an in-depth understanding of what the elements of the system were and to understand their relative importance. The team first created a list of constructs derived from the literature about the nature of the intervention—*Assertive Community Treatment* (ACT). It then interviewed research experts who understood the intervention system and could rate the importance of each feature. There was high agreement among experts about what formed the critical elements of AST. From these, the features of utmost importance were identified. These were classified as staffing, organization, or service measures of fidelity.

The research team then interviewed staff, made site visits, and examined records to assess the degree to which the criteria for fidelity had been met. In this particular case, these included measuring the client–staff ratio and the size of the clinic team as well as determining whether the team included a psychiatrist and/or a nurse, whether a primary therapist was assigned to each case, whether private offices were in a separate site from the clinic, whether the team shared cases, how often team meetings were held, whether the coordinator provided direct client service, whether there was 24-hour availability, whether time devoted to cases was limited or unlimited, whether in vivo or face-to-face contacts outside the office were conducted, whether office contacts were avoided, and the total number of client contacts. Program fidelity was correlated with mental health outcomes, with the organization and staffing subscales being significantly correlated with program outcomes (days of subsequent hospitalization). Greater adherence to the AST model thus had the benefit of possibly having greater impact on patients' mental health. The protocol and measures are available online (Teague & Bond, 2003a, 2003b).

A similar set of fidelity measurement tools are available for assessing the effectiveness of a variety of clinical services. For example, *Supported Employment Services* is designed as a vehicle for helping people with mental illnesses find employment and move toward recovery (Becker et al., 2001). *Illness Management and Recovery* (Mueser et al., 2002) is an evidence-based practice for teaching people with severe mental illness how to manage their disorder. *Family Psychoeducation* (McFarlane, Dixon, Lukens, & Lucksted, 2003) is an intervention for addressing family education and support for individuals diagnosed with schizophrenia. *Integrated Dual Disorder Treatment* (Minkoff, 1989) addresses standards for treating individuals diagnosed with both psychosis and addiction. Fidelity assessment for each of these clinical evidence-based practices includes a protocol and a measurement scale. In all of these cases, the protocol requires a rater to obtain information from agency records, interviews with employment specialists and other practitioners, supervisors, program managers, and consumers. Protocols and scales are publicly available from ebp.networkofcare.org (2012).

There are also examples of environmental interventions in schools. Unlike manual-based interventions that rely on a multisession approach, these interventions focus on the repeated application of concepts that are intended to change the climate or the culture of a school. For example, Biggs and colleagues (2008) designed a teacher self-report adherence measure for an elementary school violence prevention program, *Creating a Peaceful School Learning Environment* (CAPSLE;

Twemlow et al., 2001). The program consisted of zero tolerance for inappropriate social behavior, a plan for modeling appropriate behavior, repeated efforts to teach skills for emotional self-regulation, and a mentoring program. Thus, CAPSLE is intended as an environmental intervention in which teachers were trained and then encouraged to implement program concepts periodically during "reflection time" as well as during other times. To assess fidelity, teachers responded to four prompts that asked how often CAPSLE concepts were applied. Responses included *rarely/never, a few times per month, a few times per week, almost daily, 1–2 times per day,* and *3+ times per day*. Thus, in fidelity measurement terms, this method most closely resembles assessing dosage. This scale had high inter-item reliability (alpha = 0.84). Teachers' reports about the frequency with which program concepts were implemented were found to correlate with students' attitudes about bullying.

Teachers also reported about their attitude toward the program (intervention helpfulness) and their beliefs about their ability to influence student behavior and attitudes toward prescribed classroom management practices. Conceptually, these topics most closely reflect an assessment of quality of delivery. Two highly related questions, both assessing how helpful teachers viewed the program to be, were assessed as the "perceived value of the program" (alpha = 0.90). The three items that assessed beliefs about influence (e.g., "In your opinion, how influential are teachers for how students behave in the classroom") had acceptable internal consistency (alpha = 0.78). The six items that assessed attitudes toward the interventions' approach to classroom management had marginal internal consistency (alpha = 0.67).

Using Fidelity Outcomes

Measuring fidelity is one thing, and using these data in analyses is another. Indeed, perhaps the greatest issue facing research teams is that, once fidelity data are collected, they are rarely examined beyond providing descriptive statistics to summarize findings. This is particularly true in the dissemination environment.

There are four primary uses of fidelity data once they are collected: (1) their use as moderators or mediators in the analysis of outcome data, (2) their use as markers of sufficiency, (3) their use as feedback about performance, and (4) their use to inform the improvement of interventions. Interestingly, to date, relatively little research has used these data for any of these purposes. Each is described below, and, to the extent research has addressed these topics, summaries are presented.

Use as Mediators. The distinction between moderators and mediators was clarified by Baron and Kenny (1986). Mediators represent the mechanism through which programs achieve outcomes and, in terms of fidelity, are the immediate and intermediate effects that are hypothesized to account for how the intervention achieves its effects. Of the five types of fidelity data described above, the most common use in analysis is to use intermediate markers as mediators. These analyses are typically based on formal statistical methods first proposed by Judd and Kenny

(1981). In the case in which multiple interventions are being contrasted, simple comparative analyses may also be conducted (e.g., Hansen et al., 1991). Other researchers have contributed numerous articles that detail methods for conducting mediating variable analysis (e.g., Baron & Kenny, 1986; Hansen & McNeal, 1996; MacKinnon et al., 2007). This approach has been applied to drug prevention programs (e.g., Donaldson, Graham, & Hansen, 1994; MacKinnon et al., 1991; McNeal et al., 2004) to demonstrate how variables such as normative beliefs and commitments not to use drugs that are targeted by programs account for behavioral outcomes. These and similar analyses reveal that when programs fail to change targeted mediators that have the potential to account for behavior, programs fail (e.g., Gottfredson et al., 2009; Hansen & McNeal, 1997).

Markers of Sufficiency. As noted by Baron and Kenny (1986), moderators relate to qualities specific to the nature of the intervention itself. In terms of fidelity, the level of adherence, dosage, quality of program delivery, and level of engagement are all moderators of program delivery. Fidelity as a moderator has been examined but often post hoc, typically used to understand evaluation findings. In nearly all of these analyses, the primary focus has been on adherence as opposed to quality of delivery, dosage, or engagement. For example, Botvin and colleagues (Botvin, et al., 1995) used their assessments of proportion of objectives and main points covered (adherence) to divide intervention schools into high-fidelity and low-fidelity groups. Program effects reached significance only when more than 65 % of objectives and main points were covered. Similar effects have been seen in science education. O'Donnell (2007) found that middle school students who participated in a science intervention, the *Scaling up Curriculum for Achievement, Learning, and Equity* (SCALE-0), performed differently based on how teachers performed. In high-fidelity classrooms, students outperformed both controls and students in low-fidelity classrooms on posttest knowledge achievement tests. In this case, fidelity consisted of ratings of instructional strategies.

In the online document, "Introduction to the Evidence-Based Practice Fidelity Scales" (ebp.networkofcare.org, 2012), a key point is emphasized. Measures of fidelity when used as markers of sufficiency are best used longitudinally so that changes in fidelity can be tracked over time with the goal of marking improvement.

Feedback About Performance. An obvious use of fidelity data is to inform intervention providers about their performance. In this context, observational data may be more useful than self-report because of the potential for a positive bias. However, when self-reports are used, they may be a cause for introspection and reflection if they are correctly processed. For example, Bellg and colleagues (2004) suggest regular fidelity assessments be implemented and results be integrated into ongoing education and training of clinical staff. Central to their recommendations are assessments of adherence and dosage of treatment. These data are then used by supervisors to ensure that protocols are followed and that provider skills are developed and maintained.

Intervention Improvement. The final use of fidelity data is to improve program design and effectiveness. Donaldson (2001) outlines a general strategy for improving interventions that relies on four steps. Interventions are initially designed to

target theoretical constructs that as prior research suggests are important to change to achieve desired outcomes. The second step is then to pilot test the program and collect information about fidelity and immediate and intermediate outcomes. These data are analyzed, and if improvement is clearly possible, the intervention is revised and pilot tested again. Once the program developers are satisfied with the quality of the intervention, a full-scale implementation in a randomized control trial is conducted. However, in Donaldson's model, this is not the end. Feedback from evaluation outcomes is used to refine the intervention. Thus, in an ideal world, there will be a cycle of continuous quality improvement that will affect the content and structure of the intervention.

While relatively straightforward in concept, the method may not need to define exactly how fidelity data are used to improve the intervention. For example, in a project currently being conducted by our research team (Hansen et al., 2011), adherence (how many activities were taught), quality of delivery (4-point ratings from poor to exceptional), and quality of teaching (5-point ratings from poor to exceptional) were used to identify weak sessions. Pairs of observers coded video recordings using methods described earlier in this chapter. Once identified, these ratings and adaptations (rated from −2 to +2) were used to understand where poor teachers had consistently made errors and where exemplary teachers had made improvements to create a revised intervention, which is currently being field tested.

Conclusion

Strategies for measuring fidelity and its various aspects are diverse and often tailored to the specific needs of the researchers or the practitioners. To the extent that measures can be standardized—typically among similar types of interventions—the field can begin to move toward the systematic collection of these data. However, as the diversity of methods reviewed in this chapter attests, the needs to be met through employing fidelity measurement and the specific fidelity measures selected will often diverge. This state of affairs will not yet allow the simplification and systematization of fidelity assessment to be realized. This diversity will, in the end, lead to a common language about fidelity and a robust science for its measurement.

Among the first tasks facing the field is simplification of terms. While the terms outlined by Dane and Schneider (1998) and more recently elaborated by Durlak and DuPre (2008) provide a basic reference point, it is clear that many subtypes of each of these terms have emerged. In some cases, for instance, the challenge is to distinguish between adherence and dosage and between adherence and quality of delivery. The revision of prior publications that have confused the terms or that have provided insufficient clarity in how specific fidelity constructs have been measured is not possible. Perhaps the only reasonable strategy for moving forward is for researchers to provide evermore detail and to link their measures carefully to the traditions they are drawn from, with adequate discussion about the underlying concept that is being tapped.

In addition to such an alignment, it is also important for future research to broaden the types of measures that are employed in any one study. Few studies reviewed here collected data about all aspects of fidelity (including information about adaptation). The exemplary use of a broader range of measures should be encouraged. Further, multimodal assessments, where multiple sources of data are used, should also be encouraged. All fields of intervention will mature more quickly if these higher standards become widely adopted.

References

Abbott, R. D., O'Donnell, J., Hawkins, J. D., Hill, K. G., Kosterman, R., & Catalano, R. F. (1998). Changing teaching practices to promote achievement and bonding to school. *American Journal of Orthopsychiatry, 68*, 542–552.

Adams, A. S., Soumerai, S. B., Lomas, J., & Ross-Degnan, D. (1999). Evidence of self-report bias in assessing adherence to guidelines. *International Journal of Quality Health Care, 11*, 187–192.

Baranowski, T., Anderson, C., & Carmack, C. (1998). Mediating variable framework in physical activity interventions: How are we doing? How might we do better? *American Journal of Preventive Medicine, 15*, 266–297.

Baranowski, T., Lin, L. S., Wetter, D. W., Resnicow, K., & Hearn, M. D. (1997). Theory as mediating variables: Why aren't community interventions working as desired? *Annals of Epidemiology, 7*, S89–S95.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Basen-Enquist, K., O'Hara-Tompkins, N., Lovato, C. Y., Lewis, M. J., Parcel, G. S., & Gingiss, P. (1994). The effect of two types of teacher training on implementation of Smart Choices: A tobacco prevention curriculum. *Journal of School Health, 64*, 334–339.

Becker, D. R., Bond, G. R., McCarthy, D., Thompson, D., Xie, H., McHugo, G. J., et al. (2001). Converting day treatment centers to supported employment programs in Rhode Island. *Psychiatric Services, 52*(3), 351–357.

Bellg, A., Borrelli, B., Resnick, B., Ogedegbe, G., Hecht, J., Ernst, D., et al. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the Behavioral Change Consortium. *Health Psychology, 23*(5), 443–451.

Biggs, B., Vernberg, E., Twemlow, S., Fonagy, P., & Dill, E. (2008). Teacher adherence and its relation to teacher attitudes and student outcomes in an elementary school-based violence prevention program. *School Psychology Review, 37*(4), 533–549.

Bishop, D. C., Pankratz, M. M., Hansen, W. B., Albritton, J., Albritton, L., & Strack, J. (2013). Measuring fidelity and adaptation: Reliability of a instrument for school-based prevention programs. *Evaluation and the Health Professions*, doi: 10.1177/0163278713476882

Blakely, C. H., Mayer, J. P., Gottschalk, R. G., Schmitt, N., Davidson, W., Roitman, D. B., et al. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. *American Journal of Community Psychology, 15*, 253–268.

Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H. K. (2000). Measurement of fidelity in psychiatric rehabilitation. *Mental Health Services Research, 2*(2), 75–87.

Boruch, R. R., & Gomez, H. (1977). Sensitivity, bias, and theory in impact evaluation. *Professional Psychology, 8*, 411–433.

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., Baker, E., Dusenbury, L., Tortu, S., & Botvin, E. M. (1990). Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: Results of a 3-year study. *Journal of Consulting and Clinical Psychology, 58*, 437–446.

Botvin, G. J., Baker, E., Filazzola, A. D., & Botvin, E. M. (1990). A cognitive-behavioral approach to substance abuse prevention: One-year follow-up. *Addictive Behaviors, 15*, 47–63.

Botvin, G. J., Dusenbury, L., Baker, E., James-Ortiz, S., & Kerner, J. (1989). A skills training approach to smoking prevention among Hispanic youth. *Journal of Behavioral Medicine, 12*, 279–296.

Breitenstein, S. M., Fogg, L., Garvey, C., Hill, C., Resnick, B., & Gross, D. (2010). Measuring implementation fidelity in a community-based parenting intervention. *Nursing Research, 59* (3), 158–165.

Dane, A., & Schneider, B. H. (1998). Integrity in primary prevention programs: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23–45.

Dent, C. W., Sussman, W., Hennesy, M., Galaif, E. R., Stacy, A. W., Moss, M., et al. (1998). Implementation and process evaluation of a school-based drug abuse prevention program: Project Towards no Drug Abuse. *Journal of Drug Education, 28*(4), 361–375.

Donaldson, S. I. (2001). Mediator and moderator analysis in program development. In S. Sussman (Ed.), *Handbook of program development for health behavior research and practice* (pp. 470–496). Thousand Oaks, CA: Sage.

Donaldson, S. I., Graham, J. W., & Hansen, W. B. (1994). Testing the generalizability of intervening mechanism theories: Understanding the effects of adolescent drug use prevention interventions. *Journal of Behavioral Medicine, 17*(2), 195–216.

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology, 17*(2), 245–262.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting the implementation. *American Journal of Community Psychology, 41*, 327–350.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*, 237–256.

Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., & Falco, M. (2005). Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research, 20*, 308–313.

ebp.networkofcare.org. (2012). *Introduction to the evidence-based practice fidelity scales*. http://ebp.networkofcare.org/uploads/Fidelity_Scales_4673293.pdf. See also http://ebp.networkofcare.org/definitions/index.cfm?pageName=DefsFidelity

Ennett, S. T., Haws, S., Ringwalt, C. L., Vincus, A. A., Hanley, S., Bowling, J. M., et al. (2012). Evidence-based practice in school substance use prevention: Fidelity of implementation under real-world conditions. *Health Education Research, 26*, 361–371.

Fagan, A. A., Hanson, K., Hawkins, J. D., & Arthur, M. W. (2008). Bridging science to practice: Achieving prevention program implementation fidelity in the community youth development study. *American Journal of Community Psychology, 41*, 235–249.

Fagan, A., & Mihalic, S. (2003). Strategies for enhancing the adoption of school-based prevention programs: Lessons learned from the Blueprints for Violence Prevention replications of the LifeSkills Training Program. *Journal of Community Psychology, 31*, 235–253.

Giles, S. M., Harrington, N. G., & Fearnow-Kenney, N. (2001). Evaluation of the All Stars Program: Student and teacher factors that influence mediators of substance use. *Journal of Drug Education, 31*(4), 385–397.

Giles, S. M., Jackson-Newsom, J., Pankratz, M. M., Hansen, W. B., Ringwalt, C. L., & Dusenbury, L. (2008). Measuring quality of delivery in a substance use prevention program. *Journal of Primary Prevention, 29*, 489–501.

Gottfredson, D. C., Cross, A., Wilson, D., Rorie, M., & Connell, N. (2009). An experimental evaluation of the All Stars prevention curriculum in a community after school setting. *Prevention Science, 11*, 142–154.

Greenberg, M. T., Kusche, C. A., Cook, E. T., & Quamma, J. P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. *Development and Psychopathology, 7*, 117–136.

Hall, G. E., & Loucks, S. F. (1977). A developmental model for determining whether the treatment is actually implemented. *American Educational Research Journal, 14*, 263–276.

Hanley, S., Ringwalt, C., Vincus, A., Ennett, S., Bowling, J., Haws, S., et al. (2009). Implementing evidence-based substance use prevention curricula with fidelity: The role of teacher training. *Journal of Drug Education, 39*(1), 39–58.

Hansen, W. B. (1996). Pilot test results comparing the All Stars program with seventh grade DARE: Program integrity and mediating variable analysis. *Substance Use & Misuse, 31*, 1359–1377.

Hansen, W. B., Bishop, D. C., Dusenbury, L., Pankratz, M. M., Albritton, J., Albritton, L. P., et al. (2011). *Feeding information learned from adaptations back into program design.* Presentation made at the 20th Annual Meeting, Society for Prevention Research, Washington, DC. Retrieved from http://www.tanglewood.net/powerpoint/Hansen_SPR_2011.ppt

Hansen, W. B., Graham, J. W., Wolkenstein, B. H., Lundy, B. Z., Pearson, J. L., Flay, B. R., et al. (1988). Differential impact of three alcohol prevention curricula on hypothesized mediating variables. *Journal of Drug Education, 18*, 143–153.

Hansen, W. B., Graham, J. W., Wolkenstein, B. H., & Rohrbach, L. A. (1991). Program integrity as a moderator of prevention program effectiveness: Results for fifth grade students in the Adolescent Alcohol Prevention Trial. *Journal of Studies on Alcohol, 52*, 568–579.

Hansen, W. B., & McNeal, R. B. (1996). The law of maximum expected potential effect: Constraints placed on program effectiveness by mediator relationships. *Health Education Research, 11*, 501–507.

Hansen, W. B., & McNeal, R. B. (1997). How DARE works: An examination of program effects on mediating variables. *Health Education & Behavior, 24*, 165–176.

Hansen, W. B., Pankratz, M. M., & Bishop, D. C. (under review). Differences in observers' and teachers' assessments of fidelity and adaptation. *Health Education Research.*

Hansen, W. B., Pankratz, M. M., Dusenbury, L., Giles, S. M., Bishop, D. C., Albritton, J., Albritton, L. P., & Strack, J (2013) Styles of adaptation: The impact of frequency and valence of adaptation on preventing substance use. Health Education 113(4):345–363.

Harachi, T. W., Abbott, R. D., Catalano, R. F., Haggerty, K. P., & Fleming, C. B. (1999). Opening the black box: Using process evaluation measures to assess implementation and theory building. *American Journal of Community Psychology, 27*, 715–735.

Hill, L. G., Maucione, K., & Hood, B. (2007). A focused approach to assessing program fidelity. *Prevention Science, 8*(1), 25–34.

Johnson, K., Ogilvie, K., Collins, D., Shamblen, S., Dirks, L., Ringwalt, C., et al. (2010). Studying implementation quality of a school-based prevention curriculum in frontier Alaska: Application of video-recorded observations and expert panel judgment. *Prevention Science, 11*, 275–286.

Judd, C. M., & Kenny, D. A. (1981). *Estimating the effects of social interventions*. Cambridge, UK: Cambridge University Press.

Lillehoj, C. J., Griffin, K. W., & Spoth, R. (2004). Program provider and observer ratings of school-based preventive intervention implementation: Agreement and relation to youth outcomes. *Health Education and Behavior, 31*, 242–257.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593–614.

MacKinnon, D. A., Johnson, C. A., Pentz, M. A., Hansen, W. B., Flay, B. R., & Yang, Y. F. I. (1991). Mediating mechanisms in a school-based drug prevention program: First year effects of the Midwestern Prevention Project. *Health Psychology, 10*, 164–172.

McFarlane, W. R., Dixon, L., Lukens, E., & Lucksted, A. (2003). Family psychoeducation and schizophrenia: A review of the literature. *Journal of Marital and Family Therapy, 29*, 223–245.

McGrew, J. H., Bond, G. R., Dietzen, L. L., & Salyers, M. P. (1994). Measuring the fidelity of implementation of a mental health program model. *Journal of Consulting and Clinical Psychology, 62*, 670–678.

McNeal, R. B., & Hansen, W. B. (1999). Developmental patterns associated with the onset of drug use: Changes in postulated mediators during adolescence. *Journal of Drug Issues, 29*(2), 381–400.

McNeal, R. B., Hansen, W. B., Harrington, N. G., & Giles, S. M. (2004). How All Stars works: An examination of program effects on mediating variables. *Health Education Quarterly, 31*, 165–178.

Minkoff, K. (1989). An integrated treatment model for dual diagnosis of psychosis and addiction. *Hospital & Community Psychiatry, 40*, 1031–1036.

Mowbrey, C. T., Holter, M. C., Teague, T. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation, 24*, 315–340.

Mueser, K. T., Corrigan, P. W., Hilton, D., Tanzman, B., Schaub, A., Gingerich, S., et al. (2002). Illness management and recovery for severe mental illness: A review of the research. *Psychiatric Services, 53*, 1272–1284.

O'Donnell, C. L. (2007). *Fidelity of implementation to instructional strategies as a moderator of curriculum unit effectiveness in a large-scale middle school science quasi-experiment*. Washington, DC: George Washington University.

Pentz, M. A., Trebow, E. A., Hansen, W. B., MacKinnon, D. P., Dwyer, J. H., Johnson, C. A., et al. (1990). Effects of program implementation on adolescent drug use behavior: The Midwestern Prevention Project. *Evaluation Review, 14*, 264–289.

Ransford, C., Greenberg, M. T., Domitrovich, C. E., Small, M., & Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on implementation of a social emotional curriculum. *School Psychology Review, 38*, 510–532.

Ringwalt, C. L., Ennett, S., Johnson, R., Rohrbach, L. A., Simons-Rudolph, A., Vincus, A., et al. (2003). Factors associated with fidelity to substance use prevention curriculum guides in the nation's middle schools. *Health Education and Behavior, 30*, 375–391.

Ringwalt, C. L., Pankratz, M. M., Gottfredson, N., Jackson-Newsom, J., Dusenbury, L., Giles, S. M., et al. (2009). The effects of students' curriculum engagement, attitudes towards their teachers, and perception of their teachers' skills on school-based prevention curriculum outcomes. *Journal of Drug Education, 39*, 223–237.

Rohrbach, L. A., Dent, C. W., Skara, S., Sun, P., & Sussman, S. (2007). Fidelity of implementation in Project Towards No Drug Abuse (TND): A comparison of classroom teachers and program specialists. *Prevention Science, 8*, 125–132.

Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine, 22*, 237–260.

Rohrbach, L. A., Gunning, M., Sun, P., & Sussman, S. (2010). The Project Towards No Drug Abuse (TND) dissemination trial: Implementation fidelity and immediate outcomes. *Prevention Science, 11*, 77–88.

Ruiz-Primo, M. A. (2006). *A multi-method and multi-source approach for studying fidelity of implementation*. CSE Report 677. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing, UCLA.

Sloboda, Z., Stephens, P., Pyakuryal, A., Teasdale, B., Stephens, R. C., Hawthorne, R. D., et al. (2009). Implementation fidelity: The experience of the Adolescent Substance Abuse Prevention Study. *Health Education Research, 24*(3), 394–406.

Sobol, D. F., Rohrbach, L. A., Dent, C. W., Gleason, L., Brannon, B. R., Johnson, C. A., et al. (1989). The integrity of smoking prevention curriculum delivery. *Health Education Research, 4*, 59–67.

Spoth, R., Guyll, M., Lillehoj, C. J., Redmond, C., & Greenberg, M. (2007). PROSPER study of evidence-based intervention implementation quality by community-university partnerships. *Journal of Community Psychology, 35*, 981–999.

Spoth, R., Guyll, M., Redmond, C., Greenberg, M., & Feinberg, M. (2011). Six-year sustainability of evidence-based intervention implementation quality by community-university partnerships: The PROSPER study. *American Journal of Community Psychology, 48*(3–4), 412–425.

Spoth, R., Redmond, C., & Shin, C. (1998). Direct and indirect latent-variable parenting outcomes of two universal family-focused preventive interventions: Extending a public health-oriented research base. *Journal of Consulting and Clinical Psychology, 66*(2), 385–399.

Swanson, E., Wanzek, J., Haring, C., Ciullo, S., & McCulley, L. (2011). Intervention fidelity in special and general education research journals. *Journal of Special Education, 20*, 1–11.

Tappe, M. K., Galer-Unti, R. A., & Bailey, K. C. (1995). Long-term implementation of the Teenage Health Teaching Modules by trained teachers: A case study. *Journal of School Health, 65*.

Teague, G. B., & Bond, G. R. (2003a). *Protocol for Assertive Community Treatment Fidelity Scale*. http://ebp.networkofcare.org/uploads/ACT_Protocol_9942960.pdf

Teague, G. B., & Bond, G. R. (2003b). *Dartmouth Assertive Community Treatment Scale*. http://ebp.networkofcare.org/uploads/ACT_Scale_3113827.pdf

Teague, G. B., Bond, G. R., & Drake, R. E. (1998). Program fidelity in assertive community treatment: Development and use of a measure. *American Journal of Orthopsychiatry, 68*, 216–232.

Teasdale, B., Stephens, P. C., Sloboda, Z., Grey, S. F., & Stephens, R. C. (2009). The influence of program mediators on eleventh grade outcomes for seventh grade substance users and nonuser. *Drug and Alcohol Dependence, 102*, 11–18.

Tobler, N. S., & Stratton, H. (1997). Effectiveness of school-based drug prevention programs: A meta-analysis of the research. *Journal of Primary Prevention, 18*, 71–128.

Twemlow, S. W., Fonagy, P., Sacco, F. C., Gies, M. L., Evans, R., & Ewbank, R. (2001). Creating a peaceful school learning environment: A controlled study of an elementary school intervention to reduce violence. *American Journal of Psychiatry, 158*, 808–810.

Part III
Research Design

Chapter 16
Translating the Intervention Approach into an Appropriate Research Design: The Next-Generation Adaptive Designs for Effectiveness and Implementation Research

C. Hendricks Brown, W. Alex Mason, and Eric C. Brown

Introduction

New prevention approaches require innovative research designs to maximize our ability to obtain answers to research questions. This chapter provides an overview of new research designs that address both effectiveness and implementation research in prevention. It also includes a discussion of what we have called adaptive or hybrid designs that integrate the concepts of effectiveness and implementation into an ongoing trial. These adaptive designs have the flexibility to adapt to changes in research questions and yet maintain the rigor associated with randomized control trials. We illustrate these approaches with examples from recent studies and point the way toward the design of future trials for prevention.

Current Challenges in Prevention Intervention Research

The last two decades have seen dramatic advances in preventive intervention research due in part to carefully crafted randomized field trials that have rigorous evaluation designs. This effort has provided the field of prevention with a large and rapidly growing number of proven successful interventions for mental, emotional, and behavioral disorders. Most of these interventions have generally been tested using the classic randomized trial design for efficacy or benefit when the program is delivered under ideal circumstances. These efficacy trials involve individual-level randomization or, alternatively, group-level (e.g., school) randomization.

In contrast, testing for effectiveness, or evaluating impact under settings that are much closer to the real world (e.g., programs being delivered by practitioners in the

C.H. Brown (✉)
Department of Epidemiology and Public Health, Miller School of Medicine, University of Miami, Miami, FL, USA
e-mail: hendricks.brown@northwestern.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_16,

community rather than researchers), often requires more sophisticated designs that fit the specific needs of community and institutional settings. For example, classroom-based interventions in elementary schools can be tested with efficiency through a classroom-randomized design as long as the classes within the school are comparable (Brown & Liao, 1999). This is not often the case, as most schools use ability tracking as a way of assigning children to classes, so the balancing of classrooms within a school through random assignment of children would be required in order to achieve the higher statistical power that is possible with a classroom-based versus school-based randomized trial. By randomly assigning children to classrooms and then randomly assigning classrooms and teachers within schools to the intervention condition, i.e., using school as a blocking factor, we can minimize the effects of school and classroom variation in the design (Brown, Kellam, Ialongo, Poduska, & Ford, 2007). This is a more sophisticated statistical design than is commonly used in many efficacy studies, and it can be more challenging to conduct because it may require having schools change their criteria for classroom assignment. In addition, care must be taken to minimize leakage of the intervention into neighboring classrooms.

Even the more complex effectiveness trial design discussed above is not sufficient to answer the next sets of research questions that we face in prevention, such as which alternative implementation strategy is most beneficial for an evidence-based program. This and other research problems are discussed from the point of view of study design in this chapter. We foresee at least two important new themes that are guiding next-generation research in prevention that will require innovative designs for evaluation. First, our field is engaged in developing innovative interventions that have flexibility in how and when they are delivered, how interventions can be improved over time, and how they incorporate the subjects' own preferences. Included in this category of new types of behavioral interventions that are on the horizon are those that involve web-based (eHealth) or mobile phone (mHealth) technologies (Burns & Mohr, 2013; Mohr, Cuijpers, & Lehman, 2011; Nguyen, Cuenco, Wolpin, Benditt, & Carrieri-Kohlman, 2007; Nilsen et al., 2012). As we will see, the traditional research design is not likely to be successful in evaluating these more nimble interventions that can incorporate personal as well as contextual level needs with higher precision. Secondly, much greater attention is being paid in recent years to answering research questions around how best to implement a program, sometimes after it has achieved a status of being "evidence based" and sometimes before it has reached this evidentiary status. Addressing these new implementation or translational questions also requires innovative research designs (Spoth et al., 2013; Woolf, 2008).

In the next section we present some innovative types of designs for evaluating the effectiveness of nimble interventions, which might be Web based or utilize mobile phone technologies. Following this, we discuss three paradigms for implementation research: (1) implementing a single evidence-based intervention, (2) implementing one or more interventions selected from a menu of evidence-based interventions, and (3) implementing a program that has not yet fully demonstrated its effectiveness. These paradigms are illustrated with examples from a wide

range of prevention programs that focus on drug abuse, suicide, and HIV and across different settings including primary care, which, with the introduction of the Affordable Care Act, now provides major new prevention opportunities (Mechanic, 2012).

Evaluation of the Effectiveness of Nimble Interventions

Many of our evidence-based behavioral interventions that have undergone the most rigorous efficacy evaluations are delivered in fixed settings and given in the same sequence to all subjects. While many of these interventions have been shown to be efficacious, they are often too long, are too costly, and require too much skill on the part of the facilitator to deliver widely (Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research Advances and Promising Interventions, 2009). In addition, they tend to have limited support mechanisms to sustain them in practice (Rotheram-Borus & Duan, 2003).

Designs for Broadcasting, Narrowcasting, and Timecasting

Programs that are delivered through Web-based or cloud-based media have the opportunity to deliver effective interventions cheaply by "broadcasting" to a worldwide audience (Muñoz et al., 2006). Such widely distributed Web-delivered programs can be ideal platforms for recruiting subjects for large randomized trials. Because the marginal cost of delivering such programs to a wide audience is very small and worldwide recruitment is possible, some Web-delivered interventions can afford large samples for evaluation. Indeed, one of the advantages of a widely distributed Web-based intervention is that the intervention can be refined and tested in a continual improvement model by randomizing new individuals to an existing or a new intervention. One can afford to conduct large experiments involving upwards of a thousand subjects assigned to the existing and new intervention. With such large sample sizes, it would be possible to test for variation in impact as a function of measurable baseline characteristics (Brown et al., 2008) and ultimately deliver a version of a Web-based program that is optimal for each subgroup.

Web-based and cloud-based systems can also be directed toward a specific target audience (i.e., narrowcasting) that might be predicted to benefit directly from this particular intervention. For example, rather than screen solely for high levels of individual symptoms, a health care organization that is interested in delivering a prevention program that addresses a wide array of youth emotional and behavioral problems through parent training programs could consider screening families not for specific problem behaviors in preadolescents but rather for robust risk factors in the family that mediate intervention effects. Poor communication between parents

and children, for example, not only predicts drug abuse and HIV sexual risk behavior as well as internalizing symptoms but also leads to preventive effects when changed by a parenting intervention (Pantin et al., 2009; Prado et al., 2007, 2009, 2013; Perrino et al., Submitted for publication). By focusing on those families with poor communication, even prior to symptoms being expressed by the youth, there is an opportunity to prevent these multiple problem outcomes from occurring. One would need to test whether a parent training intervention delivered through the Web or other electronic media would be acceptable, would reach the intended audience, and would be effective. An advantage of selecting families based on this hypothesized mediator of family communication is that smaller sample sizes are required to achieve significance in such mediational designs (Brown, 1991; Pillow et al., 1991).

A new perspective on prevention that can come from considering more nimble interventions involves what can be termed "timecasting" (Brown et al., 2013). Instead of delivering a program on a set schedule, say in the evening, timecasting delivers an intervention at the time and place that is most salient. A good example involves mobile phone interventions for depression and/or suicidal behavior. Mobile phones have embedded sensors whose signals can be analyzed using machine learning to identify both patient states and contexts (Burns et al., 2011). Sensors along with ecological momentary assessment, or queries that are time sampled, can be used to identify individual patterns of behavior and to identify a state transition, for example from a low level of symptoms to a major depressive episode among those who previously had been hospitalized for a suicide attempt. Once such a state transition is identified, it could be used to intensify intervention. The design of an evaluation of such a mobile phone intervention depends on the comparison condition. If the comparison is simply traditional follow-up, say through a monthly phone call to those who had previously been hospitalized with a suicide attempt (Fleischmann et al., 2008), one could randomize individuals within each clinic, thereby using clinic as a blocking factor. However, if the key question is to determine the best circumstances for mobilizing an intervention, one could consider randomizing different intervention strategies that are planned to occur based on the occurrence of state transitions. Evaluation using such a design would need to rely on some complex modeling and causal inference. A reasonable approach would be to first develop a mixed effect bivariate growth model that models state transitions and behavioral responses under nonintervention. Then the evaluation would involve the addition of a time-dependent intervention effect. Parameters would include both immediate responses and cumulative effects from repetitive interventions.

We foresee an even greater flexibility in behavioral interventions that can be tuned to one's individual needs and lifestyle. We believe that primary care could provide an important setting for such interventions. One innovative intervention approach is to make available different options for patients accessing primary care to receive interventions, such as parent training in modalities that fit their own personal schedules and preferences. These so-called option-rich or OR intervention strategies (Pisani et al., 2013) make interventions available in ways that they would

prefer—such as face to face, Web based, or mobile phone based—rather than a single modality. By allowing each patient to identify those options that they would prefer and then randomizing each to one of these options, we can still conduct a randomized (preference) trial. One major challenge in these designs is to account for the selection preferences of different patients in making causal inferences, including those who have only one preference and cannot then be randomized as well as those who refuse to take part in a randomized trial. Two innovative approaches to this problem have been reviewed by the Cochrane Group (Preference Collaborative Review Group, 2008). One approach is to use propensity scores to investigate whether generalization of trial results to the population that is unwilling to take part in a trial is justified. The second approach is to propose two randomizations; the first is to randomize whether the subject will be in a randomized (preference) trial or a non-randomized study where only a single option is available. For those randomized to the preference arm, they would be re-randomized to one of the conditions preferred by that person (Marcus, Stuart, Shadish, & Steiner, 2012). Both of these types of designs have potential for producing valid causal inferences and predicting population-level impact better than current methods.

To date there are few trials that have used the types of designs described here, so careful evaluation of statistical power and design protocols will need to be developed. In the next section we describe innovative designs for implementation; in contrast to the above discussion there are new examples of innovative designs for implementation research, and we describe three specific examples below.

Three Implementation Paradigms

Implementation science is a young field, and researchers are only beginning to determine what scientific strategies, as well as designs, are most useful under what circumstances (Spoth et al., 2013). In this section we introduce three distinct paradigms for implementation research for prevention; we anticipate the expanded use of these designs in the future. The introductory remarks for the three paradigms then set the stage for the following subsection where one example is used to illustrate the design concerns for each paradigm.

Overview of Three Implementation Paradigms

Paradigm 1: A Head-to-Head Randomized Implementation Trial Design for a Single Evidence-Based Intervention

Commonly, it has been assumed that prior to large-scale implementation, intervention programs must first be shown to be efficacious (Flay, 1986). Efficacious

programs, then, become the best candidates for further testing in effectiveness trials (Flay et al., 2005) and, if proven effective, are prepared for widespread implementation. The first scientific paradigm focuses on such a top-down approach for wide-scale implementation of a single, effective preventive intervention. Because such an intervention is backed by strong evidence of effectiveness, there is a high likelihood that wide-scale implementation of this program, if delivered at a high level of fidelity, would lead to beneficial changes at a population level. Here the major scientific question involves how best to implement this intervention most effectively across a large spectrum of communities or organizations. As an important scientific tool to help answer this question, we consider the use of a randomized implementation trial (RIT) (Brown et al., 2009; Landsverk et al., 2012). Relatively few RITs have been used to test one implementation strategy head to head against another, holding fixed the intervention program itself. We discuss lessons learned in implementing an evidence-based treatment program for children in the foster care system requiring high levels of service. The CAL-OH randomized implementation trial, involving 51 counties in California and Ohio, serves as an example of this approach (Chamberlain, Saldana, Brown, & Leve, 2010).

Paradigm 2: Testing a Prevention Selection and Delivery System to Meet Community Needs

A second scientific paradigm to wide-scale implementation is to provide communities with a diverse set of effective prevention program options so that they can choose which of these can best meet their respective needs and available resources. In this second menu-based approach, the scientific focus is on how to support communities themselves to assess community needs, decide on which prevention programs to adopt, implement programs with fidelity, and sustain them over time. We illustrate this approach by describing the implementation of the Communities that Care (CTC) model and discuss how a randomized trial of the CTC implementation strategy is being conducted (Hawkins et al., 2008a).

Paradigm 3: Quality Improvement and Testing the Efficacy/ Effectiveness of Promising Preventive Interventions That Are Implemented in Community Settings

Our third paradigm for prevention differs dramatically from the previous two and requires more background discussion. This third scientific paradigm involves moving a promising, but not yet evidence-based, program into wider scale implementation. Here the idea is not to wait for years for the completion of rigorous effectiveness evaluation followed by a separate scientific approach toward

implementation but to combine these steps together. Glasgow and colleagues (2003) note that the characteristics of an intervention that promotes efficacy often are different from and can even run counter to those that promote effectiveness and implementation. Thus research studies that have a complementary focus on effectiveness and implementation from the onset may be fruitful in the right circumstances (Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research Advances and Promising Interventions, 2009), and consequently research designs may differ as well. The example for this third paradigm is the Common Sense Parenting (CSP) program led by Boys Town.

By no means are the three illustrative innovative, adaptive designs exhaustive. Space does not allow for more detailed presentations of other possibilities; however, other approaches are briefly mentioned in the chapter's discussion section.

Examples of Three Implementation Paradigms

Paradigm 1: A Head-to-Head Randomized Implementation Trial Design for a Single Evidence-Based Intervention: The CAL-OH Study

Program summary. Multidimensional Treatment Foster Care (MTFC) was developed in 1983 as a community-based alternative to incarceration and placing youth in residential/group care. This program is a top-tier evidence-based in-home foster care program for youth who have significant problems with chronic antisocial behavior, emotional disturbance, or delinquency. It is among the elite 11 programs that the Blueprints Project has identified as a model program (http://www.colorado.edu/cspv/blueprints/modelprograms/MTFC.html), based on multiple randomized trials showing substantial reductions in antisocial behavior, arrests and incarceration, hard drug use, and school and community problems (Chamberlain, Brown, et al., 2008; Chamberlain, Leve, & DeGarmo, 2007; Leve & Chamberlain, 2005, 2007; Leve, Chamberlain, & Reid, 2005). This program places individual youth in a family setting for 6–9 months, with foster parents recruited, selected, and trained to provide an individualized program suited to that child's needs. Daily phone calls with the foster mother are used to monitor progress; supports included a case manager, weekly supervision for the foster parent, and skill-building programs for the youth.

The traditional, individual (IND) county implementation strategy for MTFC uses the protocols developed by Treatment Foster Care Consultants, Inc. (TFCC), an agency established in 2002 to disseminate this program. An implementation expert works individually with county child welfare, juvenile justice, and mental health agencies to set up a stakeholder meeting in the county; participates in readiness planning of the county social service agencies; facilitates hiring of the

treatment team and recruitment of foster parents; and continues to monitor and support adherence and sustainability of MTFC.

Research design. In the CAL-OH study, this traditional, IND implementation strategy is tested against a Community Development Team (CDT) implementation of the same MTFC program. CDT is administered by the California Institute of Mental Health (CiMH), serving as a "broker" or an intermediary in the implementation process. The seven core processes of CDT (Sosna & Marsenich, 2006) are designed to promote motivation, engagement, commitment, persistence, and competence that lead to the development of organizational structures, policies, and procedures that support model-adherent and sustainable programs. Of particular interest is the peer-to-peer process involving leaders from multiple counties working together to solve challenges in administration, finance, and monitoring. Because CDT involves multiple counties working together, this requires that both the design and analysis of the trial account for such clustering.

The design (Chamberlain, Price, et al., 2008) of this randomized implementation trial began by determining which California counties met the following inclusion/exclusion conditions: counties could not have used MTFC in the past nor could the number of youth served in the county be too small to sustain such a program. Altogether there were 40 California counties that qualified for this program, and all agreed to participate. Because training of staff in all counties could not happen at the same time for logistical and economic reasons, we randomly assigned these counties into three cohorts of 14, 14, and 12 counties each; this assignment meant that two cohorts were wait-listed until the second or the third year to begin implementation. Within each cohort counties were also randomly assigned to the CDT or the IND implementation strategy condition. The procedures for randomization were to balance all six groupings of cohort by treatment to be as similar as possible across a wide range of county-level covariates. This was accomplished by first screening thousands of possible covariate permutations of the counties, selecting one that did an excellent job of balancing these covariates across implementation condition, and then randomly assigning each of the six groups of counties into cohort and CDT or IND. After the study began a cohort of 11 counties from Ohio were added to the study. Again these were randomly assigned to CDT or IND.

Two potential problems were anticipated in maintaining this design, as these assignments by place and time are more challenging to hold fixed compared to that of more traditional subject-randomized efficacy or effectiveness trials. In particular, not all counties assigned to the first cohort were ready to implement that first year. These few counties were permitted to wait until the next year and were replaced with a randomly selected county from the second cohort that was already assigned the same intervention condition. This replacement procedure was acceptable to the counties and did not imbalance the conditions on measured covariates.

The second anticipated challenge was acknowledging that the existing network relations between leaders of the county social service agencies in a state and CiMH brokers could potentially interfere with either the IND or the CDT condition. That is, an IND-assigned county with strong ties to CiMH colleagues could potentially

be open to receive their colleagues' advice about how to address challenges in implementation. Weekly meetings with all staffs, including CiMH brokers, were held to maintain advice channels as per the study protocol. A related concern was that IND- and CDT-assigned counties would seek each others' advice about problems in implementation. A subsequent social network analysis, however, revealed that there were very few influence network connections that directly crossed these two conditions (Palinkas et al., 2013).

Comparison of implementation success across the two strategies was done by examining the speed, quality, and quantity with which implementation took place. The Stages of Implementation Completion (SIC) was the primary outcome measure that provides this information at the county level (Chamberlain, Brown, & Saldana, 2011). The analytic methods used are similar to those used to predict the time it took for counties to begin adoption (Brown, Chamberlain, Saldana & Wang, under review; Saldana, Chamberlain, Wang, & Brown, 2011; Wang, Saldana, Brown, & Chamberlain, 2010).

Strengths and limitations of the design. This type of design could well be used to test other implementation strategies against one another, holding fixed the single evidence-based intervention. Such head-to-head RITs that are tested in separate trials with different evidence-based interventions could be used to assess the general utility of an implementation "scaffold," i.e., a strategy like CDT that should be applicable to implementing a wide range of specific programs. RITs can also be used to test the degree to which adaptations in the intervention itself should be permitted or controlled (Brown et al., 2008, 2009; Landsverk et al., 2012). Thus one implementation arm would be more permissive of adaptations than the other, yet would still provide a framework of what kinds of adaptations are permitted. This RIT design could be useful in addressing whether standardized programs can be implemented effectively in minority populations or those with health disparities or whether such prevention programs should have adaptations that specifically address these communities' needs (Brown et al., 2013). Because this RIT requires that a preselected intervention be used, one appropriate use of this design is when a particular program's use is mandated, such as was true in California Child Welfare system, or is likely to be adopted by many if not most communities. Otherwise too many communities may need to be excluded from this design. As with all designs, RITs are good at answering the primary research questions but have limited ability to answer others. In the CAL-OH study, there is no random assignment of youth to MTFC; counties are able to choose which youth should be provided these services. These unobservable selection factors may account for differences in youth outcomes between those who were provided MTFC and those who were not. Thus there is limited ability in this particular RIT design to examine questions of effectiveness on youth outcomes. The use of RITs in other settings may also have limited ability to assess the impact on the target population's health outcomes unless random selection of subjects or propensity score adjustments are used. In addition, to date RIT designs are relatively expensive to carry out, and this may prohibit their use in some settings.

Paradigm 2: Testing a Prevention Selection and Delivery System to Meet Community Needs: The Community Youth Development Study

Program summary. The availability of a large number of effective prevention programs and strategies for mental, emotional, and behavioral problems in youth prompts a call for identifying effective prevention selection and delivery systems that can successfully implement these programs and strategies in communities. A systems-based approach to delivering prevention programs and strategies suggests that the underlying infrastructure of a community, which comprises the relevant service sectors and constituencies of the community, can be oriented in a collaborative manner toward a mutually agreed upon outcome. This orientation would require a transformation of a community's "existing policies and practices, resource allocations, relational structures, community norms and values, and skill and attitudes" (Foster-Fishman & Behrens, 2007, p. 192). As noted in the literature on systems theory, successful system transformation should be guided by a clearly articulated purpose and theory of change, explicate the interdependencies among components, be *proactive* as opposed to *reactive*, focus on measurement and needs assessment, and incorporate feedback loops and processes for system realignment (Behrens & Foster-Fishman, 2007; Fixsen et al., 2005; Spoth & Greenberg, 2011; Wandersman et al., 2008; Weisz et al. 1995). These perspectives are integrated into the CTC prevention selection and delivery system described below and tested in the Community Youth Development Study (CYDS) randomized implementation trial.

CTC is a prevention strategy that takes a systems-based approach to engage community coalitions in selecting and implementing appropriate prevention programs that have been tested and found to be effective (Hawkins et al. 2008b). CTC incorporates a *community diagnostic approach* (Feinberg et al., 2012) to prevention by identifying elevated levels of youth risk and depressed levels of youth protective factors via epidemiologic assessments of risk and protective factors for adolescent health and behavior problems. These assessments are conducted through school-based administrations of the Communities That Care Youth Survey (Arthur et al., 2002), which contains valid and reliable measures of risk/protective factors and outcomes identified in community, school, family, and peer–individual domains.

CTC incorporates elements of implementation and dissemination in three distinct areas. First, CTC catalyzes communities for prevention action through a series of six trainings delivered by certified trainers to community leaders, key stakeholders, and coalition board members during a 6–12-month CTC implementation phase. Second, communities enter a prevention system transformation phase where, at a macro systems level, communities are theorized to increase (a) the adoption of science-based approaches to prevention, (b) collaboration among relevant community sectors, (c) community norms against adolescent drug use, (d) community support for prevention initiatives, and (e) use of *social development strategies* (Hawkins et al., 2010) for positive youth development. Third, based on epidemiologic data collected in their respective communities, community coalition members

develop community action plans, in which appropriate tested-effective prevention programs are selected from a list of evidence-based programs and are implemented. Program fidelity is monitored and assessed by the community coalition to ensure maximal impact on targeted youth outcomes. These areas of implementation and dissemination represent leverage points for change at macro (i.e., broad community change), meso (i.e., coalition functioning), and micro (i.e., program delivery to and uptake by youth and families) levels.

Research design. The CYDS is a community-randomized controlled trial of the CTC prevention system tested against communities that are not given such supports (Brown et al., 2009; Hawkins et al. 2008b). The CYDS consists of 24 communities in seven states across the United States that were matched into pairs with one community from each matched pair randomized to receive the CTC system or a "prevention-as-usual" comparison condition. To date, CYDS communities have been followed over years with repeated surveys of community leaders, prevention coalition members, and both cross-sectional and panel samples of youth. Although the primary focus of the CYDS was to assess CTC's impact on youth risk, protection, and antisocial behavior outcomes (e.g., drug use, delinquency, and violence), examination of the implementation and dissemination processes making up CTC's theory of change has demonstrated successful impact of CTC at multiple levels of prevention system development and transformation.

For example, Quinby et al. (2008) examined the fidelity of CTC implementation in the CYDS using the CTC Milestones and Benchmarks Tool with coalition coordinators and found that the CTC system could achieve the desired number of benchmarks within the designated time frame and be completely implemented with fidelity within an average of 11 months. Moreover, 11 of the 12 CTC coalitions continued to exist over 6 years later (20 months after support to the communities ended) and the proportion of benchmarks met by CTC coalitions did not decline significantly during this period (Arthur et al., 2010; Gloppen et al., 2012). Significant macro-level changes in prevention systems were reported at 18 months, 4.5 years, and 6.5 years after CTC installation (Brown et al., 2007; 2011). Compared with control communities, communities implementing CTC exhibited greater levels of adopting and sustaining a science-based approach to prevention and greater growth in community norms against adolescent drug use during the three follow-up periods than did control communities. Finally, examinations of program fidelity indicated significantly greater numbers of tested and effective programs and strategies implemented in CTC communities compared to control communities (Fagan et al., 2012). Also, high levels of adherence to program implementation protocols, quality of program delivery, participant responsiveness, and intervention dosage and exposure to the prevention programs and strategies that were selected in CTC communities (2004–2008) were obtained and maintained throughout the course of the study (Fagan et al., 2008, 2011, 2012).

Significant intervention effects on youth reduced levels of risk factors and incidence and prevalence of adolescent alcohol, cigarette, smokeless tobacco, binge drinking, and delinquent behaviors (Hawkins et al., 2008a, 2009) in the CYDS panel sample. Cost-effectiveness analyses due to reduced incidence and

prevalence of antisocial behaviors indicate that $5.30 in savings are realized in long-term benefits for every dollar invested in CTC (Kuklinski et al., 2012).

Strengths and limitations of the design. A design that randomizes communities to a prevention system like CTC or one with no support allows one to assess the overall system's impact on youth behavior. It can also test hypothesized mediators of the implementation process itself. Testing a broad prevention system like CTC that does not preselect a single preventive intervention is much more natural for most communities than a prevention approach that is restricted to a single intervention as discussed above. Allowing communities at least some choice in selecting programs and/or focus age groups or outcomes should enhance community participation in such trials. Such a design works well in small communities in rural areas and likely not as well in large urban areas where multiple community coalitions are likely to exist. Also, some restrictions in the choice of population (e.g., age group) are required in order to evaluate the impact on comparable youth outcomes. What the CYDS design did not address and cannot answer is whether the delivery system of training and support for communities to implement CTC could be improved. Such questions of comparing two different implementation strategies, for example, would require a head-to-head RIT design much like that of the CAL-OH project described above.

Paradigm 3: Quality Improvement and Testing the Efficacy/ Effectiveness of Promising Preventive Interventions That Are Implemented in Community Settings: Common Sense Parenting

Program summary. The CSP study involves a recently initiated randomized prevention trial of the CSP parent-training program (Burke, Herron, & Barnes, 2006). CSP was developed at Boys Town, a large and well-established service provider with an integrated continuum of care ranging from intensive residential treatment to community support programs. CSP is a promising and widely used preventive intervention, annually serving more than 1,500 children from 1,000 families at Boys Town sites and also disseminated to other agencies across 47 states and 14 countries.

Common Sense Parenting® (CSP) was adapted from the *Boys Town Family Home Program*, which itself is an adaptation of the Teaching Family Model (Phillips et al., 1973; Wolf et al., 1976) that applied behavioral treatment interventions in a group home setting for abused, neglected, and behavior-problem youths (Davis & Daly, 2003). The theoretical basis for CSP draws from social learning principles (Bandura, 1977) as well as from social interaction theory and coercion theory (Snyder et al., 1994). The program has several components that are found in evidence-based parenting interventions (Barth et al., 2005). Specifically, CSP promotes parenting skills that encourage positive behavior, discourage negative

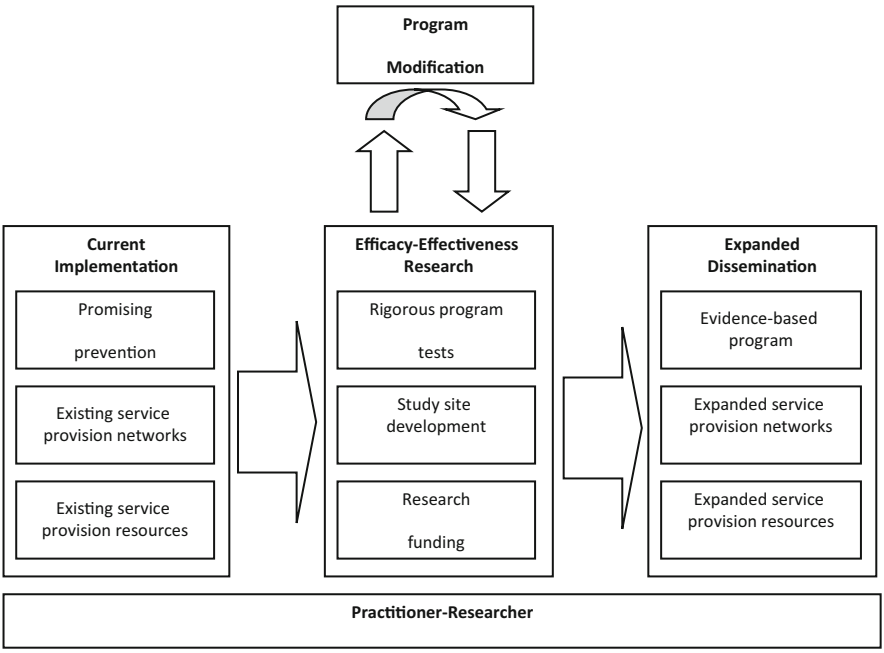

Fig. 16.1 A framework for improving and testing the efficacy/effectiveness of promising programs already being implemented

behavior, and teach alternatives to problem behavior in children. In a series of six-weekly, 2-h group workshops, parents learn and practice skills that address issues of communication, discipline, decision making, relationships, self-control, and school success.

CSP has a program manual in the form of a trainer's guide that walks workshop leaders through the content and activities of each session (Burke, Schuchmann, & Barnes, 2006), along with a parenting book for participants (Burke, Herron, & Barnes, 2006) and manuals for a 3-day CSP Training-of-Trainer's workshop. Typically, some modification of material is needed to facilitate evaluation. For example, although implementation fidelity forms had been created for CSP, they were designed primarily for supervision rather than research and had to be adapted for the randomized trial.

Research design. The three-step research strategy for implementing as well as improving the quality of CSP is illustrated in Fig. 16.1. This model begins with a promising prevention program that appears to be beneficial in its original form as currently being implemented (see left-hand column of Fig. 16.1). Boys Town conducted a series of CSP evaluation studies throughout the 1990s across a variety of samples and settings. Results indicated statistically and clinically significant improvements in child behavior, parent attitudes, family satisfaction, and parent problem-solving ability from pretest to posttest (Thompson et al., 1997) and at

3-month follow-up (Thompson et al., 1993). Pretest–posttest designs are not sufficient to demonstrate effectiveness, which is the second step in Fig. 16.1.

Such promising programs are then modified based on effectiveness research (middle column of Fig. 16.1). The standards of evidence increasingly require rigorous designs such as RCTs, regression discontinuity designs, and quasi-experimental studies with propensity score matching or similar techniques for handling potential selection biases (Shadish et al., 2002). The current effectiveness study of CSP is a randomized trial. We discuss several key considerations in evaluating CSP for effectiveness.

The first challenge was obtaining the support of key service provision stakeholders for committing to a program test and accepting the ensuing results. Program developers and service providers were reluctant to subject their interventions to such testing while these programs were being implemented in specific sites. Researchers successfully appealed to the desire of most service providers to achieve maximum positive impact with their program by using a rigorous test to know with confidence that a program works as designed.

A second challenge was the selection of sites. One option that maximizes external validity is to design a study that draws on existing service networks to target individuals who are already served by the intervention as part of routine service delivery. However, when sampling from existing service provision networks the existing pool of targeted program participants may be too small to provide a sufficient sample size for a rigorous prevention trial and may place constraints on randomization. Alternatively, a new, independent sample can be targeted for recruitment into a prevention trial. Developing new study sites may reduce external validity but may provide more flexibility in terms of study design and research protocols.

Balancing these types of considerations, this RCT is recruiting a new sample in the CSP trial, independent of the routine program delivery provided by Boys Town. Families of eighth-grade students enrolled in selected middle schools in the Tacoma Public School District comprise the pool of eligible participants. This decision provides maximum flexibility in implementing random assignment to conditions and places the study site in proximity to research partners in the Social Development Research Group, University of Washington, which is leading the data collection effort for the trial.

Research has confirmed that combining youth training with parent training can provide added value (Demarsh & Kumpfer, 1986; Kazdin et al., 1992), especially for families with adolescent-aged children. With this knowledge, this trial evaluated the standard-form CSP against a modified form that adds youth involvement and skill training components to the original curriculum. The modified CSP program, known as CSP Plus, is an eight-session program that adds materials from the Stepping Up To High School curriculum, developed at SDRG as a booster session for the Raising Health Children Project (Catalano et al., 2003). In the trial, recruited families are randomly assigned to a standard CSP program condition, a modified CSP Plus program condition, or a no-intervention control condition to test the efficacy of each program and determine the added value of the enhancements.

The ultimate goal of the three-stage framework in Fig. 16.1 is to promote expanded implementation (Mason et al., 2013). The existing implementation

supports for the original program should be naturally supportive of the revised intervention as well, thus speeding implementation once a program is found to be effective. Also, success of an intervention as measured by an effectiveness trial can open the doors to service provision funding opportunities supporting access to new clients through expanded service provision networks. Primarily, this occurs through the exposure and marketing opportunities that derive from having a supported program listed on evidence-based program registries, as many service provision funding opportunities now require selecting programs that are listed as evidence based.

Strengths and limitations of the design. As noted, there are numerous prevention programs being implemented in communities throughout the United States, many of which have never been evaluated (Spoth, 2008). A large number of these programs likely are without benefit or may even cause harm. Given the array of programs available, the significant resources required to conduct the efficacy and effectiveness research, and the high costs of investing in ineffectual or harmful programs, it is critical to identify promising programs that hold strong potential for being proven efficacious and effective. In this regard, there are at least three considerations. First, a promising prevention program should have characteristics found in evidence-based preventive interventions (Embry & Biglan, 2008; Nation et al., 2003; Small, Cooney, & O'Connor, 2009) and should be based on theory with a logic model whose core components, active ingredients, or "kernels" are found in proven successful preventive interventions (e.g., Embry & Biglan, 2008). Although the underlying theories and corresponding logic models of promising programs that are being implemented in applied community settings may not be as well articulated as those of programs created in academic research contexts, they need to be consistent with current prevention science (Nation et al., 2003).

A second characteristic of promising prevention programs is that the program must be evaluable (Rutman, 1980; Smith, 1989). An essential requirement for evaluation is that the program be manualized and well documented (Small et al., 2009; West, 2008); monitoring tools for documenting the fidelity of implementation are also required to ensure that the intervention is delivered as designed (Dane & Schneider, 1998).

Promising programs also need preliminary evidence for their success. Many service providers collect short-term process and outcome data on clients who complete their programs, which can be useful for tracking changes in participants over time. Subject to the validity threats inherent in non-experimental data (Shadish, Cook, & Campbell, 2002), positive changes are consistent with the possibility of intended program effects. Small-scale evaluations, typically using single-group, pre-post, or even quasi-experimental designs, can provide other evidence. Identifying, improving, and testing promising programs already in use throughout communities may provide important public health benefit. In some areas of prevention, notably suicide prevention (Brown, Wyman, Guo, & Peña, 2006; Wyman et al., 2008, 2010), very few programs are known to be efficacious or effective. In these situations, it makes sense to select promising programs that already have wide usage, as some of the major implementation challenges have already been worked out through trial and error. These programs can then be

prepared for rigorous effectiveness testing through the development of well-specified intervention manuals and assessment tools to examine hypothesized mediating pathways. Further discussion of this approach is presented in the Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research Advances and Promising Interventions' Institute of Medicine report on prevention (Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research Advances and Promising Interventions, 2009).

It is also possible to use this type of design strategy for an intervention that was found to be ineffective but whose prevention delivery system is worth using. The Drug Abuse Resistance Education (D.A.R.E.) program, for example, has been one of the most popular and widely used substance use prevention programs in the United States (Ringwalt et al., 2002), and it has been used in up to 80 % of our communities (Hallfors & Godette, 2002), yet results from several reviews of this program as it has been delivered have concluded that the program in its initial form is ineffective in preventing substance use (Ennett, Tobler, Ringwalt, & Flewelling, 1994; Lynam et al., 1999; West & O'Neal, 2004). Given the popularity of D.A.R.E. and the extensive national delivery system infrastructure it already has in place across the country through D.A.R.E. America, the program has been recognized as a potential candidate for enhancements. Indeed, there have been two major effectiveness tests of a modified program built around D.A.R.E. One, known as D.A.R.E. Plus, has revealed some improvements, but positive intervention effects of D.A.R.E. Plus are qualified (e.g., restricted to boys; Perry et al., 2003) and have been limited in scope (Eischens, Komro, Perry, & Bosma, 2004). Another approach expanding on the same opportunity of a highly functioning prevention delivery system was the Adolescent Substance Abuse Prevention Study (ASAPS), a large randomized trial supported by the Robert Wood Johnson Foundation and led by Dr. Zili Sloboda. This provided a new curriculum called Take Charge of Your Life (TCYL) that resembled Life Skills Training but was delivered like the original D.A.R.E. program by police officers with the support of the national D.A.R.E. America training network. This program delivered in the 7th and 9th grades showed overall moderate negative impact on adolescent alcohol and tobacco use at the time of the 11th-grade follow-up for those who were not using these substances when in 7th grade but with statistically significant benefit for marijuana use among those who used marijuana when in the 7th grade (Sloboda et al., 2009).

Discussion

We have described a wide range of design options to address two major themes that we suggest will become increasingly important for prevention. Behavioral interventions are expected to become more nimble, allowing more flexibility in delivery times, populations, and modalities. We foresee that interventions that are option rich will require much more complex designs that take into account individual

variation, time, and context. The second major theme for prevention is a focus on implementation questions rather than effectiveness per se. We illustrated the use of a rollout randomized implementation trial to evaluate a single evidence-based intervention. Additionally randomized implementation trials can be used to evaluate a broad-based prevention delivery program such as CTC. We also consider randomized trial evaluations of implementation and quality improvement for promising programs already being implemented in communities but requiring more rigorous effectiveness evaluation and perhaps revision of the program itself.

There are actually a wide range of opportunities to link effectiveness and implementation designs together in the same study. The traditional model of research is to separate studies into small distinct research agendas ranging from pre-intervention, to efficacy, to effectiveness, and to implementation (Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research Advances and Promising Interventions, 2009; Landsverk et al., 2012). However, to accelerate the research agenda of moving from effectiveness to wide-scale implementation we need to use more intelligent, adaptive designs. The concept of an adaptive design originated in pharmaceutical trials where joint phase I and II adaptive designs answer toxicity and appropriate doses and phase II/III trials bridge the gap between dosage and efficacy. These and other ideas for adaptive designs have now been developed with specific application to public health (Brown et al., 2009). There are numerous opportunities for preventive interventions to use adaptive designs that address multiple stages of research. One place is in a single 5-year experiment that has one specific aim to test effectiveness and a second to test one or more implementation research questions. An example of this approach is demonstrated in a recent classroom-based intervention trial (Poduska et al., 2009). The Whole-Day program was designed to integrate a behavioral and reading curriculum seamlessly into first-grade classrooms. Since this was a classroom-based intervention, children were randomized to first-grade classes and these classes were in turn randomized to intervention or control conditions in the first year. Coaches were used to support the delivery of the Whole-Day intervention to these randomly assigned classrooms. In a follow-up year the scalability of this intervention, i.e., an implementation question, was tested by extending the intervention to all the classrooms. Because the number of coaches remained the same as the first year but twice as many classrooms were involved, this scalability year tested the impact when the coaching level was reduced.

We suggest one important opportunity for an adaptive design that could be used to accelerate the implementation of an intervention that is found to be highly effective for a serious outcome such as prevention of AIDS among HIV-infected individuals. Despite the availability of efficacious antiviral treatments, some marginalized populations with HIV have difficulty maintaining the high levels of adherence that is needed to achieve benefit. Text reminders (SMS) delivered technically through the support of a research group (i.e., an efficacy study) appear to be helpful for certain low adhering populations at least based on the results of a small pilot study (Dowshen, Kuhns, Johnson, Holoyda, & Garofalo, 2012; Harris et al., 2013; Lewis et al., 2013). In designing a full-scale randomized efficacy trial

of this text-based intervention, one may want to consider and plan for the possibility of a highly successful increase in adherence, an impact so large that it could necessitate the early stopping of a trial because of ethical concerns. But merely finding out that the intervention is efficacious and offering these text messages to those in the control condition are not sufficient to address the next major challenges as to how community-based organizations or other service agencies would be able to implement and sustain such an effective intervention. Rather than have the study be terminated and then an entirely new grant proposal be developed to answer next sets of implementation questions, it would be more efficient to transition from efficacy to implementation within the same research project. Provided there is adequate planning for an eventuality of a highly beneficial intervention, this approach could lead to a much more rapid integration of successful interventions. We recommend that this potential be recognized explicitly in the protection of human subject section of research grants on randomized trials that offer such potential for making major impacts in addressing the critical needs for treatment for AIDS and other major health issues. With prior planning in the grant application and support from the Data Safety and Monitoring Committee and the funding agency, it would be possible to shift a grant that has a highly successful intervention to a rigorous implementation study without losing momentum.

As the field of prevention science radically adapts in investigating different strategies for addressing challenging local, national, and international prevention needs, innovative research designs are desired that help clarify which programs and implementation strategies are most useful and under what conditions. The use of these innovative designs requires methodological development and evaluation of their statistical properties. But just as important in designing such studies are the partnerships between policy makers, community leaders, private partners, and researchers in bringing these crucial interventions into practice (Brown et al., 2012; Kellam, 2012; Poduska et al., 2012).

Acknowledgments We thank all our research colleagues and California and Ohio participants on the CAL-OH project, the Center for Prevention Implementation Methodology (Ce-PIM), and the Implementation Methods Research Group (IMRG). For her work in leading the CAL-OH project we thank Dr. Patricia Chamberlain. For the CYDS, the authors would like to acknowledge Dr. J. David Hawkins, Principal Investigator of the CYDS and Founding Director of the Social Development Research Group, School of Social Work, University of Washington. We gratefully acknowledge our colleagues on the CSP prevention trial and colleagues and the 24 communities in the CYDS trial. We especially thank our funders: National Institute on Drug Abuse (NIDA) for funding Ce-PIM (P30DA027828) along with the Office of Behavioral and Social Science Research, NIDA finding of the CYDS project (R01DA025651, R01DA015183) with co-funding from the National Cancer Institute, the National Institute of Child Health and Human Development, the National Institute of Mental Health, the Center for Substance Abuse Prevention, and the National Institute on Alcohol Abuse and Alcoholism. We also acknowledge support from the National Institute of Mental Health for the Prevention Science and Methodology Group (R01MH040859), the CAL-OH project (R01MH076158), and the IMRG Center (P30MH074678). We additionally thank the National Institute of General Medical Sciences (NIGMS) and the National Institute of Minority Health and Health Disparities (NIMHD) for support under the University of Miami's Center for Translational Science (UL1TR000460) and the National Institute of Allergy and Infectious Diseases (NIAID) Miami Center for AIDS Research

(P30AI073961). The content of this chapter is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

Arthur, M. W., Hawkins, J. D., Pollard, J. A., Catalano, R. F., & Baglioni, A. (2002). Measuring risk and protective factors for use, delinquency, and other adolescent problem behaviors the communities that care youth survey. *Evaluation Review, 26*(6), 575–601.

Arthur, M. A., Hawkins, J. D., Brown, E. C., Briney, J. S., & Oesterle, S. (2010). Implementation of the communities that care prevention system by coalitions in the community youth coalition study. *Journal of Community Psychology, 38*(2), 245–258.

Bandura, A. (1977). *Social learning theory*. Englewoods Cliffs, NJ: Prentice-Hall.

Barth, R. P., Landsverk, J., Chamberlain, P., Reid, J. B., Rolls, J. A., Hurlburt, M. S., et al. (2005). Parent-training programs in child welfare services: Planning for a more evidence-based approach to serving biological parents. *Research on Social Work Practice, 15*(5), 353–371.

Behrens, T. R., & Foster-Fishman, P. G. (2007). Developing operating principles for systems change. *American Journal of Community Psychology, 39*(3–4), 411–414.

Brown, C. H. (1991). Comparison of mediational selected strategies and sequential designs for preventive trials: Comments on a proposal by pillow et al. *American Journal of Community Psychology, 19*(6), 837–846.

Brown, C. H., Chamberlain, P., Saldana, L., & Wang, W. (under review). Evaluation of two implementation strategies in fifty-one counties in two states: Results of a randomized implementation trial. Submitted to *Implementation Science*.

Brown, C. H., Kellam, S. G., Ialongo, N., Poduska, J., & Ford, C. (2007). Prevention of aggressive behavior through middle school using a first grade classroom-based intervention. In M. T. Tsuang, M. J. Lyons, & W. S. Stone (Eds.), *Towards prevention and early intervention of major mental and substance abuse disorders* (pp. 347–370). Arlington, VA: American Psychiatric Publishing, Inc.

Brown, E. C., Hawkins, J. D., Arthur, M. W., Briney, J. S., & Fagan, A. A. (2011). Prevention service system transformation using communities that care. *Journal of Community Psychology, 39*(2), 183–201. doi:10.1002/jcop.20426.

Brown, C. H., Kellam, S. G., Kaupert, S., Muthén, B. O., Wang, W., Muthén, L., ... McManus, J. (2012). Partnerships for the design, conduct, and analysis of effectiveness, and implementation research: Experiences of the prevention science and methodology group. *Administration & Policy in Mental Health, 39*(4), 301–316. doi: 10.1007/s10488-011-0387-3.

Brown, C. H., & Liao, J. (1999). Principles for designing randomized preventive trials in mental health: An emerging developmental epidemiology paradigm. *American Journal of Community Psychology, 27*(5), 673–710.

Brown, C. H., Mohr, D. C., Gallo, C. G., Mader, C., Palinkas, L. A., Wingwood, G., ... Jacombs, C. (2013). A computational future for preventing HIV in minority communities: How advanced technology can improve implementation of effective programs. *Journal of Acquired Immune Deficiency Syndrome, 63*(Suppl 1), S72–S84.

Brown, C. H., Ten Have, T. R., Jo, B., Dagne, G., Wyman, P. A., Muthen, B., et al. (2009). Adaptive designs for randomized trials in public health. *Annual Review of Public Health, 30*, 1–25.

Brown, C. H., Wang, W., Kellam, S. G., Muthén, B. O., Petras, H., Toyinbo, P., ... The Prevention Science and Methodology Group (2008). Methods for testing theory and evaluating impact in randomized field trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug and Alcohol Dependence, 95*(Suppl 1), S74–S104; Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.drugalcdep.2008.01.005. doi: http://dx.doi.org/10.1016/j.drugalcdep.2007.11.013

Brown, C. H., Wyman, P. A., Guo, J., & Peña, J. (2006). Dynamic wait-listed designs for randomized trials: New designs for prevention of youth suicide. *Clinical Trials, 3*(3), 259–271.

Burke, R., Herron, R., & Barnes, B. A. (2006). *Common sense parenting: Using your head as well as your heart to raise school-aged Children* (3rd ed.). Boys Town, NE: Boys Town Press.

Burke, R. V., Schuchmann, L. F., & Barnes, B. A. (2006). *Common sense parenting: Trainer guide* (3rd ed.). Boys Town, NE: Boys Town Press.

Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., et al. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research, 13*(3), e55. doi:10.2196/jmir.1838.

Burns, M. N., & Mohr, D. C. (2013). eHealth, telehealth, and telemedicine. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of behavioral medicine*. New York, NY: Springer.

Catalano, R. F., Mazza, J. J., Harachi, T. W., Abbott, R. D., Haggerty, K. P., & Fleming, C. B. (2003). Raising healthy children through enhancing social development in elementary school: Results after 1.5 years. *Journal of School Psychology, 41*(2), 143–164.

Chamberlain, P., Brown, C. H., & Saldana, L. (2011). Observational measure of implementation progress: The stages of implementation completion (SIC). *Implementation Science, 6*(1), 116–124. doi:10.1186/1748-5908-6-116.

Chamberlain, P., Brown, C., Saldana, L., Reid, J., Wang, W., Marsenich, L., … Bouwman, G. (2008). Engaging and recruiting counties in an experiment on implementing evidence-based practice in California. *Administration and Policy in Mental Health and Mental Health Services Research, 35*, 250–260.

Chamberlain, P., Leve, L. D., & DeGarmo, D. S. (2007). Multidimensional treatment foster care for girls in the juvenile justice system: 2-year follow-up of a randomized clinical trial. *Journal of Consulting and Clinical Psychology, 75*(1), 187–193. http://dx.doi.org/10.1037/0022-006X.75.1.187.

Chamberlain, P., Price, J., Leve, L. D., Laurent, H., Landsverk, J. A., & Reid, J. B. (2008). Prevention of behavior problems for children in foster care: Outcomes and mediation effects. *Prevention Science, 9*(1), 17–27.

Chamberlain, P., Saldana, L., Brown, C. H., & Leve, L. D. (2010). Implementation of multidimensional treatment foster care in California: A randomized trial of an evidence-based practice. In M. Roberts-DeGennaro & S. Fogel (Eds.), *Empirically supported interventions for community and organizational change* (pp. 218–234). Chicago, IL: Lyceum Books, Inc.

Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research Advances and Promising Interventions. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: Nation Academies Press.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23–45.

Davis, J. L., & Daly, D. L. (2003). *Girls and boys town long-term residential program training manual* (4th ed.). Boys Town, NE: Father Flanagan's Boys' Home.

DeMarsh, J., & Kumpfer, K. L. (1986). Family-oriented interventions for the prevention of chemical dependency in children and adolescents. *Journal of Children in Contemporary Society, 18*(1–2), 117–151.

Dowshen, N., Kuhns, L. M., Johnson, A., Holoyda, B. J., & Garofalo, R. (2012). Improving adherence to antiretroviral therapy for youth living with HIV/AIDS: A pilot study using personalized, interactive, daily text message reminders. *Journal of Medical Internet Research, 14*(2), e51. doi:10.2196/jmir.2015.

Eischens, A., Komro, K. A., Perry, C. L., Bosma, L. M., & Farbakhsh, K. (2004). The association of extracurricular activity participation with substance use among youth in the DARE plus project. *American Journal of Health Education, 35*(2), 68–75.

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*(3), 75–113. doi:10.1007/s10567-008-0036-x.

Ennett, S. T., Tobler, N. S., Ringwalt, C. L., & Flewelling, R. L. (1994). How effective is drug abuse resistance education? A meta-analysis of project DARE outcome evaluations. *American Journal of Public Health, 84*(9), 1394–401.

Fagan, A. A., Arthur, M. W., Hanson, K., Briney, J. S., & Hawkins, J. D. (2011). Effects of communities that care on the adoption and implementation fidelity of evidence-based prevention programs in communities: Results from a randomized controlled trial. *Prevention Science, 12*(3), 223–234.

Fagan, A. A., Hanson, K., Briney, J. S., & Hawkins, J. D. (2012). Sustaining the utilization and high quality implementation of tested and effective prevention programs using the communities that care prevention system. *American Journal of Community Psychology, 49*(3–4), 365–377.

Fagan, A., Hanson, K., Hawkins, J., & Arthur, M. (2008). Bridging science to practice: Achieving prevention program implementation fidelity in the community youth development study. *American Journal of Community Psychology, 41*(3), 235–249. doi:10.1007/s10464-008-9176-x.

Feinberg, M., Jones, D., Cleveland, M., & Greenberg, M. (2012). The community epidemiology of underage drinking: Variation across communities in relations of risk to alcohol use. *Prevention Science, 13*(6), 551–561. doi:10.1007/s11121-012-0281-6.

Feinberg, M., Greenberg, M., Osgood, D., Sartorius, J., & Bontempo, D. (2007). Effects of the communities that care model in Pennsylvania on youth risk and problem behaviors. *Prevention Science, 8*(4), 261–270. doi:10.1007/s11121-007-0073-6.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. (No. FMHI Publication #231). Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute.

Foster-Fishman, P. G., & Behrens, T. R. (2007). Systems change reborn: Rethinking our theories, methods, and efforts in human services reform and community-based change. *American Journal of Community Psychology, 39*(3–4), 191–196.

Flay, B. R. (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine, 15*(5), 451–474.

Flay, B., Biglan, A., Boruch, R., Castro, F., Gottfredson, D., Kellam, S., ... Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*(3), 151–175. doi: 10.1007/s11121-005-5553-y.

Fleischmann, A., Bertolote, J. M., Wasserman, D., De Leo, D., Bolhari, J., Botega, N. J., ... Thanh, H. T. T. (2008). Effectiveness of brief intervention and contact for suicide attempters: A randomized controlled trial in five countries. *Bulletin of the World Health Organization, 86*(9), 703–709.

Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health, 93*(8), 1261–1267.

Gloppen, K. M., Arthur, M. W., Hawkins, J. D., & Shapiro, V. B. (2012). Sustainability of the communities that care prevention system by coalitions participating in the community youth development study. *The Journal of Adolescent Health, 51*(3), 259–264.

Hallfors, D., & Godette, D. (2002). Will thePrinciples of effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research, 17*(4), 461–470.

Harris, J. L., Furberg, R., Martin, N., Kuhns, L., Lewis, M. A., Coomes, C., ... Uhrig, J. D. (2013). Implementing an SMS-based intervention for persons living with human immunodeficiency virus. *Journal of Public Health Management and Practice, 19*(2), E9–E16. doi: 10.1097/PHH.0b013e3182582b59; 10.1097/PHH.0b013e3182582b59.

Hawkins, J. D., Shapiro, V. B., & Fagan, A. A. (2010). Disseminating effective community prevention practices: Opportunities for social work education. *Research on Social Work Practice, 20*(5), 518–527. doi:10.1177/1049731509359919.

Hawkins, J. D., Oesterle, S., Brown, E. C., Arthur, M. W., Abbott, R. D., Fagan, A. A., et al. (2009). Results of a type 2 translational research trial to prevent adolescent drug use and delinquency: A test of communities that care. *Archives of Pediatrics & Adolescent Medicine, 163*(9), 789.

Hawkins, J., Catalano, R., Arthur, M., Egan, E., Brown, E., Abbott, R., et al. (2008a). Testing communities that care: The rationale, design and behavioral baseline equivalence of the community youth development study. *Prevention Science, 9*(3), 178–190. doi:10.1007/s11121-008-0092-y.

Hawkins, J. D., Brown, E. C., Oesterle, S., Arthur, M. W., Abbott, R. D., & Catalano, R. F. (2008b). Early effects of communities that care on targeted risks and initiation of delinquent behavior and substance use. *Journal of Adolescent Health, 43*(1), 15–22.

Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. *Journal of Consulting and Clinical Psychology, 60*(5), 733–747. doi:10.1037/0022-006X.60.5.733.

Kellam, S. (2012). Developing and maintaining partnerships as the foundation of implementation and implementation science: Reflections over a half century. *Administration and Policy in Mental Health, 39*(4), 317–320. doi:10.1007/s10488-011-0402-8.

Kuklinski, M. R., Briney, J. S., Hawkins, J. D., & Catalano, R. F. (2012). Cost-benefit analysis of communities that care outcomes at eighth grade. *Prevention Science, 13*(2), 150–161.

Landsverk, J., Brown, C. H., Chamberlain, P., Palinkas, L., Rolls Reutz, J., & Horwitz, S. M. (2012). Design and analysis in dissemination and implementation research. In R. C. Brownson, G. A. Colditz, & E. K. Proctor (Eds.), *Dissemination and implementation research in health: Translating science to practice*. London: Oxford University Press.

Leve, L. D., & Chamberlain, P. (2005). Association with delinquent peers: Intervention effects for youth in the juvenile justice system. *Journal of Abnormal Child Psychology, 33*(3), 339–347. http://dx.doi.org/10.1007/s10802-005-3571-7.

Leve, L., & Chamberlain, P. (2007). A randomized evaluation of multidimensional treatment foster care: Effects on school attendance and homework completion in juvenile justice girls. *Research on Social Work Practice, 17*, 657–663.

Leve, L. D., Chamberlain, P., & Reid, J. B. (2005). Intervention outcomes for girls referred from juvenile justice: Effects on delinquency. *Journal of Consulting & Clinical Psychology, 73*(6), 1181–1185.

Lewis, M. A., Uhrig, J. D., Bann, C. M., Harris, J. L., Furberg, R. D., Coomes, C., et al. (2013). Tailored text messaging intervention for HIV adherence: A proof-of-concept study. *Health Psychology, 32*(3), 248–253. doi:10.1037/a0028109.

Lynam, D. R., Milich, R., Zimmerman, R., Novak, S. P., Logan, T. K., Martin, C., et al. (1999). Project DARE: No effects at 10-year follow-up. *Journal of Consulting and Clinical Psychology, 67*(4), 590–593. doi:10.1037/0022-006X.67.4.590.

Marcus, S. M., Stuart, E. A., Shadish, W. R., & Steiner, P. M. (2012). Estimating the causal effect of randomization versus treatment preference in a doubly-randomized preference trial. *Psychological Methods, 17*(2), 244–254.

Mason, W. A., Fleming, C. B., Thompson, R. W., Haggerty, K. P., & Snyder, J. J. (2013). A framework for testing and promoting the expanded dissemination of promising preventive interventions that are being implemented in community settings. Prevention Science.

Mechanic, D. (2012). Seizing opportunities under the affordable care act for transforming the mental and behavioral health system. *Health Affairs, 31*(2), 376–382. doi:10.1377/hlthaff.2011.0623.

Mohr, D. C., Cuijpers, P., & Lehman, K. (2011). Supportive accountability: A model for providing human support to enhance adherence to eHealth interventions. *Journal of Medical Internet Research, 13*(1), e30. doi:10.2196/jmir.1602.

Muñoz, R. F., Lenert, L. L., Delucchi, K., Stoddard, J., Perez, J. E., Penilla, C., et al. (2006). Toward evidence-based internet interventions: A Spanish/English web site for international smoking cessation trials. *Nicotine & Tobacco Research, 8*(1), 77–87.

Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., et al. (2003). What works in prevention: Principles of effective prevention programs. *The American Psychologist, 58*(6–7), 449–456. doi:10.1037/0003-066X.58.6-7.449.

Nguyen, H. Q., Cuenco, D., Wolpin, S., Benditt, J., & Carrieri-Kohlman, V. (2007). Methodological considerations in evaluating EHealth interventions. *Canadian Journal of Nursing Research, 39*(1), 116–134.

Nilsen, W., Kumar, S., Shar, A., Varoquiers, C., Wiley, T., Riley, W. T., ... Atienza, A. A. (2012). Advancing the science of mHealth. *Journal of Health Communication, 17*, 5–10. doi: 10.1080/10810730.2012.677394.

Palinkas, L. A., Holloway, I. W., Rice, E., Brown, C. H., Valente, T., & Chamberlain, P. (2013). Influence network linkages across implementation strategy conditions in a randomized controlled trial of two strategies for scaling up evidence-based practices in public youth-serving systems. *Implementation Science, 8*, 133.

Pantin, H., Prado, G., Lopez, B., Huang, S., Tapia, M. I., Schwartz, S. J., ... Branchini, J. (2009). A randomized controlled trial of Familias unidas for Hispanic adolescents with behavior problems. *Psychosomatic Medicine, 71*(9), 987–995. doi: 10.1097/PSY.0b013e3181bb2913.

Perrino, T., Pantin, H., Prado, G., Huang, S., Howe, G., Beardslee, W., et al. (Submitted for publication). Preventing internalizing symptoms among hispanic adolescents: A synthesis across Familias unidas trials. *Prevention Science*

Perry, C. L., Komro, K. A., Veblen-Mortenson, S., et al. (2003). A randomized controlled trial of the middle and junior high school D.A.R.E. and D.A.R.E. plus programs. *Archives of Pediatrics & Adolescent Medicine, 157*(2), 178–184. doi:10.1001/archpedi.157.2.178.

Phillips, E. L., Phillips, E. A., Wolf, M. M., & Fixsen, D. L. (1973). Achievement place: Development of the elected manager system. *Journal of Applied Behavior Analysis, 6*(4), 541–561.

Pillow, D. R., Sandler, I. N., Braver, S. L., Wolchik, S. A., Gersten, J. C., Pillow, D. R., ... Gersten, J. C. (1991). Theory-based screening for prevention: Focusing on mediating processes in children of divorce. *American Journal of Community Psychology, 19*(6), 809–836.

Pisani, A. R., Wyman, P. A., Petrova, M., Schmeelk-Cone, K., Goldston, D. B., Xia, Y., et al. (2013). Emotion regulation difficulties, youth-adult relationships, and suicide attempts among high school students in underserved communities. *Journal of Youth and Adolescence, 42*(6), 807–820.

Poduska, J., Gomez, M., Capo, Z., & Holmes, V. (2012). Developing a collaboration with the Houston independent school district: Testing the generalizability of a partnership model. *Administration and Policy in Mental Health and Mental Health Services Research, 39*(4), 258–267. doi:10.1007/s10488-011-0383-7.

Poduska, J., Kellam, S. G., Brown, C. H., Ford, C., Windham, A., Keegan, N., et al. (2009). Study protocol for a group randomized controlled trial of a classroom-based intervention aimed at preventing early risk factors for drug abuse: Integrating effectiveness and implementation research. *Implementation Science, 4*, 56. doi:10.1186/1748-5908-4-56.

Prado, G., Huang, S., Cordova, D., Malcolm, S., Estrada, Y., Cano, N., ... Brown, C. H. (2013). Ecodevelopmental and intrapersonal moderators of a family based preventive intervention for Hispanic youth: A latent profile analysis. *Prevention Science*. doi: 10.1007/s11121-012-0326-x.

Prado, G., Pantin, H., Briones, E., Schwartz, S. J., Feaster, D., Huang, S., ... Szapocznik, J. (2007). A randomized controlled trial of a parent-centered intervention in preventing substance use and HIV risk behaviors in Hispanic adolescents. *Journal of Consulting and Clinical Psychology, 75*(6), 914–926. doi: 10.1037/0022-006X.75.6.914.

Prado, G. J., Schwartz, S. J., Maldonado-Molina, M., Huang, S., Pantin, H. M., Lopez, B., et al. (2009). Ecodevelopmental x intrapersonal risk: Substance use and sexual behavior in Hispanic adolescents. *Health Education & Behavior, 36*(1), 45–61.

Preference Collaborative Review Group. (2008). Patients' preferences within randomised trials: Systematic review and patient level meta-analysis. *British Medical Journal*. doi: 10.1136/bmj.a1864

Quinby, R. K., Hanson, K., Brooke-Weiss, B., Arthur, M. W., Hawkins, J. D., Hanson, K., et al. (2008). Installing the communities that care prevention system: Implementation progress and fidelity in a randomized controlled trial. *Journal of Community Psychology, 36*(3), 313–332.

Ringwalt, C. L., Ennett, S., Vincus, A., Thorne, J., Rohrbach, L. A., & Simons-Rudolph, A. (2002). The prevalence of effective substance use prevention curricula in U.S. middle schools. *Prevention Science, 3*(4), 257–265. doi:10.1023/A:1020872424136.

Rotheram-Borus, M. J., & Duan, N. (2003). Next generation of preventive interventions. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(5), 518–526.

Rutman, L. (1980). *Planning useful evaluations: Evaluability assessments*. Thousand Oaks, CA: Sage Publications.

Saldana, L., Chamberlain, P., Wang, W., & Brown, C. H. (2011). Predicting program start-up using the stages of implementation measure. *Administration and Policy in Mental Health, 39* (6), 419–425. doi:10.1007/s10488-011-0363-y.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton, Mifflin, and Company.

Sloboda, Z., Stephens, R. C., Stephens, P. C., Grey, S. F., Teasdale, G., Hawthorne, R. D., et al. (2009). The Adolescent Substance Abuse Prevention Study: A Randomized Field Trial of a Universal Substance Abuse Preventive Intervention. *Drug and Alcohol Dependence, 102*, 1–10.

Small, S. A., Cooney, S. M., & O'Connor, C. (2009). Evidence-informed program improvement: Using principles of effectiveness to enhance the quality and impact of family-based prevention programs. *Family Relations, 58*(1), 1–13.

Smith, M. F. (1989). *Evaluability assessment: A practical approach*. Boston, MA: Kluwer Academic Publishers.

Snyder, J., Edwards, P., McGraw, K., Kilgore, K., & Holton, A. (1994). Escalation and reinforcement in mother-child infant conflict: Social processes associated with the development of physical aggression. *Development and Psychopathology, 6*, 305–321.

Sosna, T., & Marsenich, L. (2006). *Community development team model: Supporting the model adherent implementation of programs and practices*. Sacramento, CA: California Institute for Mental Health.

Spoth, R. (2008). Translating family-focused prevention science into effective practice: Toward a translational impact paradigm. *Current Directions in Psychological Science, 17*(6), 415–421.

Spoth, R., & Greenberg, M. (2011). Impact challenges in community science-with-practice: Lessons from PROSPER on transformative practitioner-scientist partnerships and prevention infrastructure development. *American Journal of Community Psychology, 48*(1–2), 106–119.

Spoth, R., Rohrbach, L. A., Greenberg, M., Leaf, P., Brown, C. H., Fagan, A., ... Society for Prevention Research Type 2 Translational Task Force Members and Contributing Authors. (2013). Addressing core challenges for the next generation of type 2 translation research and systems: The translation science to population impact (TSci impact) framework. *Prevention Science*. doi: 10.1007/s11121-012-0362-6.

Thompson, R. W., Grow, C. R., Ruma, P. R., Daly, D. L., & Burke, R. V. (1993). Evaluation of a practical parenting program with middle- and low-income families. *Family Relations, 42*(1), 21–25.

Thompson, R. W., Ruma, P. R., Brewster, A. L., Besetsney, L. K., & Burke, R. V. (1997). Evaluation of an air force child physical abuse prevention project using the reliable change index. *Journal of Child and Family Studies, 6*(4), 421–434. doi:10.1023/A:1025093328618.

Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., et al. (2008). Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. *American Journal of Community Psychology, 41*(3–4), 171–181. http://dx.doi.org/10.1007/s10464-008-9174-z

Wang, W., Saldana, L., Brown, C. H., & Chamberlain, P. (2010). Factors that influenced county system leaders to implement an evidence-based program: A baseline survey within a randomized controlled trial. *Implementation Science, 5*(1), 72.

Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*(5), 688–701.

West, R. (2008). Providing full manuals and intervention descriptions: Addiction policy. *Addiction, 103*, 1141.

West, S. L., & O'Neal, K. K. (2004). Project D.A.R.E. outcome effectiveness revisited. *American Journal of Public Health, 94*(6), 1027–1029. doi:10.2105/AJPH.94.6.1027.

Wolf, M. M., Phillips, E. L., Fixsen, D. L., Braukmann, C. J., Kirigin, K. A., Willner, A. G., et al. (1976). Achievement place: The teaching-family model. *Child and Youth Care Forum, 5*(2), 92–103. doi:10.1007/BF01555232.

Woolf, S. (2008). The meaning of translational research and why it matters. *Journal of the American Medical Association, 299*, 211–213.

Wyman, P. A., Brown, C. H., Inman, J., Cross, W., Schmeelk-Cone, K., Guo, J., et al. (2008). Randomized trial of a gatekeeper training program for suicide prevention: Impact on school staff after one year. *Journal of Consulting and Clinical Psychology, 76*(1), 104–115.

Wyman, P. A., Brown, C. H., Lomurray, M., Schmeelk-Cone, K., Petrova, M., Yu, Q., et al. (2010). An outcome evaluation of the Sources of Strength suicide prevention program delivered by adolescent peer leaders in high schools. *American Journal of Public Health, 100* (9), 1653–1661.

Chapter 17
The Epidemiologic Case-Crossover and Case–Control Approaches in Prevention Research

James C. (Jim) Anthony

Introduction

Point: If the evidence from prevention research is to be definitive, compelling, and translational into public health action, our studies must be tightly controlled and disciplined, perhaps even rigid. For this reason, within the prevention sciences, we tend to bite the tongue when our ideas track toward innovation in research approach.

Counterpoint: It can be argued that Professor Janine Jagger of Virginia Commonwealth University is the most highly decorated prevention scientist of the early twenty-first century. She won the 2002 MacArthur "genius" award competition and holds multiple patents for her innovations in prevention research approaches and their impact on public health in the domain of occupational exposures to blood-borne pathogens. How many readers of this volume know of her innovations and work? My own guess: not many.

Analysis: Analyzed abstractly, from a distance, the Jagger approach might be described as an imaginative and innovative application of basic principles learned through careful study of injury epidemiology and injury control as taught by Susan Baker, William Haddon, and other pioneers in that specific domain of public health research. Expressed in very concrete terms, the principles include working backward from the injurious event (exposure to, or infection by, a blood-borne pathogenic agent) to its most proximal antecedents—before turning attention to larger macro-level influences (e.g., legislation). To illustrate, the injury epidemiologist who interviews a hospital worker just found to be infected by a blood-borne pathogen (e.g., via an errant needle) might ask, "What were you doing just before the injury?" or "What action were you taking just before that happened?" or "What

J.C. Anthony (✉)
Epidemiology & Biostatistics, Michigan State University, 909 Fee Road, MSU campus, East Lansing, MI 48824, USA
e-mail: janthony@msu.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_17,

were you doing just before the injury that was different from what you had done before?" or "What device were you using?"

In a pendulum swing back to a more abstract level of thinking, we might characterize this approach as crunching the time dimension toward a limit of zero, where "zero time" is the moment when the injurious event has occurred. Working from that moment, we can take very thin slices of time as we work backward from the event toward the most proximal determining influences.

For most prevention researchers in the behavioral and social sciences, with a strong lifespan developmental tradition, these retrospective thoughts might be alien, given the orientation of lifespan developmental research to the more distal influences expressed within the ecodevelopmental framework from the moment of conception onward. Retrospective thinking also might be alien to the behavior analyst accustomed to studying how the consequences and what comes after a behavior actually might exert a governing influence on whether the behavior is repeated.

Nonetheless, retrospective thinking of this type is part and parcel of what epidemiologists learn as they seek mastery over threats to validity in their empirical studies. Whereas a key facet of behavior analysis is crunching the time dimension toward a limit of zero as we look at the consequences of a behavior, a key facet of epidemiologic analysis is crunching the time dimension toward a limit of zero as we look at the most proximal antecedents of a behavior, experience, injurious event, onset of a pathological process, or disease state transition.

The epidemiologic case-crossover design, introduced more than 20 years ago by Maclure (1990), can be distilled in relation to that simple question as might be asked by a clinician on seeing a patient in acute pain or distress: "What were you doing just before [the pain] that was different from what you had done before?" The beauty and power of this particular epidemiologic research approach derive from its subject-as-own-control constraints on potentially confounding antecedent variables, irrespective of whether they have been measured.

This chapter provides a very basic introduction, for prevention researchers, to the epidemiologic case-crossover design, which generally is neglected in reviews of prevention science methods. The chapter also covers potential utility, for prevention research, of the more traditional between-subject designs: (1) the nested case–control design with incidence density sampling and (2) the classical case–control design. The case–cohort design is described briefly, in passing, mainly to help readers draw distinctions among the several varieties of sampling approaches in epidemiologic case–control research.

Epidemiologic Case-Crossover Research

Introducing the concept of the within-subject or subject-as-own-control epidemiologic case-crossover research approach a little more than 20 years ago, Maclure (1990) drew attention to what many call "the third variable" problem in

observational studies (i.e., confounding variables that, if left uncontrolled, can lead to incorrect inferences about cause–effect relations). He noted that many variables in the domain of confounding are within-subject individual-level characteristics of a relatively stable character. In prevention research, one might think of the genetic or cultural architecture of one's family heritage, the family trajectory of socioeconomic position, or more distal variables such as personality traits. In theory, when prevention research involves randomization, the distribution of confounding variables of this type should be brought into balance by randomization into study subgroups (e.g., intervention and control conditions). Nonetheless, when the between-group comparisons of such variables are made, there can be imbalances and residual confounding, even when the randomization procedures have been implemented with fidelity. Alternatively, during the course of the preventive trial, there can be differential sample attrition or measurement distortions that are determined, at least in part, by these confounding variables such that originally balanced distributions of the variable become imbalanced for reasons other than the effect of the preventive intervention.

Maclure's innovative epidemiologic case-crossover approach can be nested within the prevention research design or can be used as a stand-alone research approach prior to randomized experiments, to hold in check the more stable family history, personality trait, and other within-subject confounding variables. In brief, as is true for all epidemiologic research on acute conditions, the point of departure for the case-crossover approach is a predefined population with some of the population members experiencing acute onsets of a disease, injury, or other adverse health events, while others are spared. Without predefinition of a population or theoretical specification of a population from which these newly incident "cases" arise, it is difficult if not impossible to characterize the study an epidemiologic study. Nonetheless, this condition is readily satisfied in most rigorous prevention research, where randomized assignments or related maneuvers sort population members into intervention and "control" or "comparison" or "standard-setting" subgroups.

Once the population has been predefined or specified, the epidemiologic case-crossover research approach focuses on the occurrence of the acute conditions we might seek to prevent. For illustration, think of a car crash or the first heart attack or the first panic attack or the first use of cocaine. Each preventable acute-onset condition has been studied using the epidemiologic case-crossover approach, with some resulting evidence that is more definitive than the evidence from prior between-subject studies.

In the public health approach to car crashes, the prevailing hypothesis is that use of a mobile telephone while driving might account for excess risk of a crash. All sorts of individual-level confounding variables might be posited to explain away any observed association in between-subject comparisons of drivers with and without car crashes and their use of mobile phones. For example, drivers who choose to use their phones while driving might be less harm avoidant and more reckless than drivers who choose not to do so; they might drive more recklessly. Using the epidemiologic case-crossover approach with its within-subject or subject-

as-own-control features, Redelmeier and Tibshirani (1997) were able to constrain these individual-level personality traits and predisposing characteristics; the empirical estimates point toward quadrupling of risk of these car crashes attributable to use of cellular phones while driving, when these sources of variation are held constant. The study was carried out by querying the cellular phone records of almost 700 drivers who had just crashed and by discovering that the odds of cellular phone use during a defined interval before the driver's crash were greater than the odds of cellular phone use during a prior control interval when no crash had occurred.

Perspicacious prevention scientists will appreciate that we are unlikely ever to be able to make randomized assignments of mobile phone usage to some drivers while others are denied mobile phone usage. They also will appreciate that the primary sources of potential confounding in research of this type can be traced back to that set of individual-level personality traits and predispositions that are relatively stable over time; these are precisely the sources of variation that are held constant in the within-subject or subject-as-own-control designs. It follows that we have opened the door to the possibility that the epidemiologic case–control design might be useful in prevention research—to an extent that is generally greater than has been appreciated to this date.

The architecture of epidemiologic case-crossover research, after predefinition or specification of the population from which newly incident events or cases arise, involves designation of a hazard interval during the life of the affected case, typically just before the acute onset of the event of interest, as well as a control interval of roughly equal duration, most often sampled from an interval before the hazard interval, as depicted in Fig. 17.1 (The hazard and control intervals are lined up because all subjects had the outcome event occurring at the same point in time. This is not an essential element in the case-crossover approach (see Gordis, 2009; Maclure, 1990; Szklo & Nieto, 2007). All temporally stable within-subgroup sources of variation are held constant in this type of case-crossover approach, and the evidence supporting the alternative to the null hypothesis about a time-varying causal variable ($X = 1$) is found in the number of subjects with $X = 1$ in the hazard interval but $X \sim = 1$ in the control interval, divided by the number of subjects with $X = 1$ in the control interval with $X \sim = 1$ in the hazard interval).

In some instances, a range of potential control intervals are studied, as in O'Brien, Comment, Liang, and Anthony (2012), whose case-crossover research focused on the possibility that, in the United States, the onset of cannabis smoking might trigger onset of cocaine use, perhaps because entry into the cannabis market in the United States brings the user into greater proximity with friends (or vendors) who also have cocaine on hand.

Wu and Anthony (2000), as well as O'Brien et al. (2012), provide a general introduction to the basic elements of the epidemiologic case-crossover research design and sketch its pertinence in a variety of research contexts, including prevention research. For this chapter, it may be most helpful to add that the outcomes of prevention research often are conceptualized in relation to non-acute processes, such as the development of a pattern of maladaptive behavior or the gradual

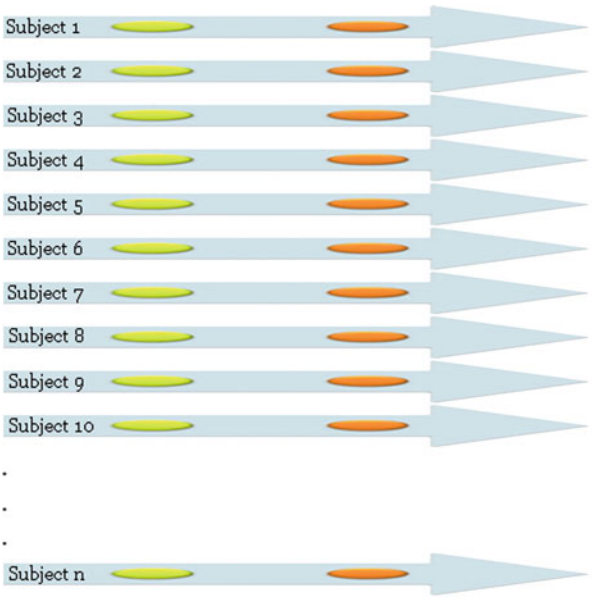

Fig. 17.1 Epidemiologic case-crossover approach, with each subject's control interval (*yellow*) and each subject's hazard interval (*orange*), as time passes

emergence of a state of psychological distress or some other departure from a state of psychological well-being. The epidemiologic case-crossover research approach is not particularly well suited to studies of these "insidious onset" conditions. The approach is more pertinent to discrete state transitions or acute onsets (e.g., car crash, first panic attack, first use of cocaine). Nevertheless, even in the study of the more insidious onset conditions, there often can be found state transitions or acute onsets of public health significance, such as the first suicide ideation or attempt in the progression toward a diagnosable major depressive episode or the first chance to try a drug in the progression toward a drug dependence syndrome. Prevention scientists, with imagination, can harness the epidemiologic case-crossover approach to their advantage when the pathological processes under study qualify as "punctuated disequilibrium" processes with state transitions and acute onsets. Increasing use of ecological momentary assessments (EMAs) (or experience sampling methods) in large sample prevention research will bring new opportunities for this epidemiologic case-crossover research approach.

For prevention scientists interested in the epidemiologic case-crossover approach, Gordis (2009) provides a basic introduction; Szklo and Nieto (2007) provide additional details. A basic introduction and applications in the context of biobehavioral research are supplied by Wu and Anthony (2000) and by O'Brien et al. (2012). One of the most recent and informative examples pertinent to the prevention sciences has been contributed by Warner et al. (2005), who used the case-crossover approach to try to rule out unmeasured confounding variables in their study of the degree to which condom usage might be preventing sexually transmitted infections.

The research potential of the epidemiologic case-crossover approach nested within ongoing prevention trials and long-span developmental studies has been neglected to this point—possibly because prevention scientists are not fully aware of the ease with which this approach can be nested in these designs. In most instances, the process of becoming a success or failure in relation to a prevention research outcome can be seen in terms of one or more acute state transitions that occur during the outcome process. Examples already mentioned include a child's first chance to try a drug in the process of increasing drug involvement, the first panic attack in the process of developing panic disorder, and the first discrete suicide ideation event. The generalizable "dynamic cascade" model introduced by Dodge and colleagues (2009) lends itself to the identification of these acute state transitions as part of each "cascade" in the process. By applying the innovations introduced by Jagger and Maclure, prevention scientists can add new elements to their research designs and will gain a more thorough understanding of the determinants of success and failure of their preventive interventions.

Nested Case–Control Research with Incidence Density Sampling

The epidemiologic case-crossover approach crunches the time dimension toward a limit of zero in its focus on an individual's odds of an exposure during a hazard interval just before the acute crash, injury, or disease event of interest, relative to the individual's odds of that exposure during some well-selected control interval, all estimated within the subject-as-own-control design. The epidemiologic approach known as case–control research with incidence density sampling is a between-subject design that crunches the time dimension toward a limit of zero in sampling of newly incident cases and non-cases during follow-up of a predefined or specified epidemiologic population or sample. That is, a population or population sample free of disease (or adverse health event) is followed forward in time. As each newly incident case is detected, there is a sampling of those still "at risk" of becoming cases, and thereafter there is a comparison of the observed odds of a suspected causal exposure value in the newly incident cases relative to the expected odds of that suspected causal exposure value among the non-case controls (to that instantaneous time point). In this description, the between-individual character of the nested case–control research design should be clearly distinguishable from the within-subject or subject-as-own-control design described above in the section on the epidemiologic case-crossover approach. The nestedness of this case–control research design refers to its position within the context of the ongoing life experiences of a predefined or specified epidemiologic population (e.g., inhabitants of an epidemiologic catchment area or health district) or within the context of an ongoing prospective longitudinal study, which might originate as a randomized prevention trial. The incidence density sampling refers to the fact that, as each newly incident

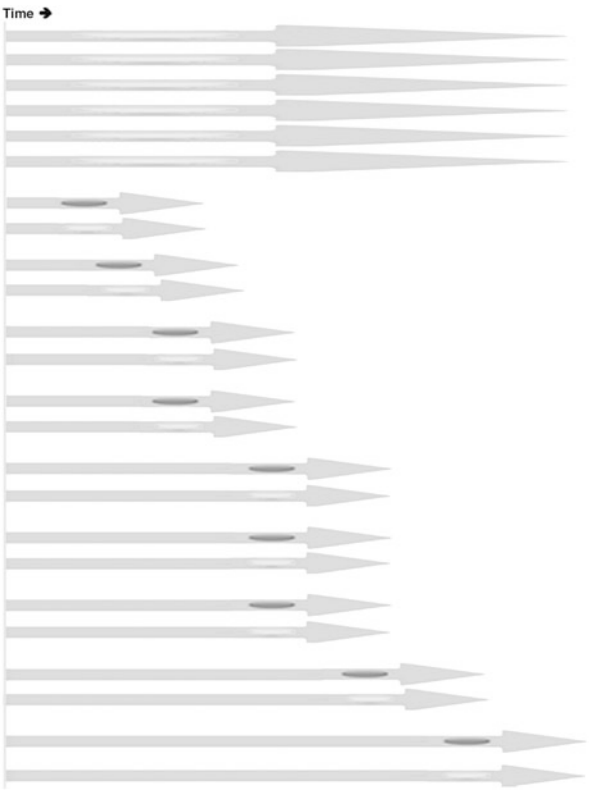

Fig. 17.2 Epidemiologic case–control approach with incidence density sampling; each between-subject comparison lines up the time of diagnosis of the newly incident case with the time of sampling one or more time-matched non-case controls

case is identified and there is review of the odds of exposure at or before that time point, there also is a concurrent time sampling of the population under study to identify corresponding non-case controls, with review of the odds of exposure at or before that time point in the life of the control, as depicted in Fig. 17.2 (Most individuals survive the follow-up interval and do not become diagnosed as newly incident cases (top rows of long arrows, colored yellow to indicate eligibility to be time sampled as non-case controls throughout the study follow-up interval). As time passes and newly incident cases start to be diagnosed, each newly incident case (orange) is matched in time with one or more non-case controls. Evidence supporting the alternative to the null hypothesis about a time-varying causal variable ($X = 1$) is found in the degree to which the odds of seeing newly incident cases with $X = 1$ before the outcome condition exceeds the odds of seeing a time-matched non-case control with $X = 1$ before that same time. In contrast with the case-crossover approach shown in Fig. 20.1, temporally stable within-subject sources of variation are not necessarily held constant in this type of between-subject case–control comparison (due to the between-subject design). Nor is there any guarantee that time-varying sources of variation are being held constant, which depends on other variables that might be matched at the time of sampling non-case

controls (e.g., gender, age) or that might be measured and held constant (e.g., via multiple logistic regression or other causal modeling approaches).

It is fair to ask what this type of nested case–control design has to do with prevention research. Several illustrations may be useful, starting with the topical and contemporary interpretation in genetic mutations and other genetic susceptibility traits pertinent as influences on distal targets of prevention research. To illustrate, if there is ongoing prevention research that attempts to reduce the risk of autism, one suspects that parental age has been measured for all population members, in which case a nested case–control design with incidence density sampling is unnecessary; the exposure variable is measured for all members of the population sample. Nonetheless, consider the possibility that there has been no assessment, at baseline or in prior assessments, of the SCN2A locus on chromosome 2 (sodium channel, voltage-gated, type II, alpha subunit), recently identified with replication as a genetic susceptibility trait for autism (and in prior studies for infantile febrile seizures). Whereas the identification of this potentially important genetic influence on autism traits is recent (Sanders et al., 2012), there is no reason for any ongoing prevention research teams to throw up their hands and confess that they have not measured this characteristic. Instead, precisely because the genetic susceptibility trait now is conceptualized as a time-invariant trait, it can be measured in the newly incident autism cases of the prevention research study, as they are detected, and it also can be measured in a relatively sample of non-case controls followed to the point of detection of each newly incident case (where the non-case controls have not experienced newly incident autism by the time of detection of the case). Here, the crunching of the time dimension involves the slice of time when the newly incident case is detected, and the goal is to secure assessments of that newly incident case (e.g., serum or saliva specimens) as soon as the case is detected, with corresponding bioassays for a modest-sized sample of non-case controls who have survived to the point of case detection but who have not (yet) become cases.

The mention of a modest-sized sample of non-case controls may come as a surprise to some prevention researchers accustomed to thinking of every study enterprise as one that must be applied to all members of the population sample. On occasion, there may be reasons to treat everyone exactly the same (e.g., concern about measurement reactivity or assay intervention interactions), but often this is not the case. If so, it actually is inefficient to draw more than 8–10 non-case controls at the time each newly incident case is detected; in many instances, the asymptotes for relative statistical power and precision of study estimates are reached with as few as 4–6 non-case controls per case (Szklo & Nieto, 2007). This is good news for a prevention researcher in the early stages of long-term follow-up, during which newly incident cases are being detected quite slowly, as illustrated in the Kellam and Anthony (1998) randomized trial to examine the possible preventive effect of the Good Behavior Game with respect to onset of smoking during middle childhood and early adolescence.

The intersection of EMAs and experience sampling methods (e.g., via smartphone push technology) with ongoing prevention research will create more

opportunities for the use of this form of nested case–control study with incidence density sampling, with between-subject designs complementing within-subject case-crossover research nested within the same study. The smartphone approach makes it possible to gain time samples of the experience of study participants throughout a prevention trial's follow-up process, efficiently and cheaply. One difficulty is management of the resulting high-dimensionality data when it comes time to make use of the EMA trace. Section 17.2 sketched how an EMA trace might be studied quite readily via the case-crossover approach, with a restricted view of the newly incident cases, time-sampled during sets of "hazard" and "control" intervals. There also can be time sampling of newly incident cases and non-case controls with incidence density sampling so that the entire mass of EMA data become manageable. That is, using the between-subject nested case–control approach with incidence density sampling, the research team would detect a newly incident case and would look back to one or more EMA assay results just before the occurrence of case status (i.e., the "hazard" interval), as well as to one or more EMA assay results for sampled "control" intervals in the life of that person. This within-subject contrast provides the case-crossover evidence for or against causal inference. Concurrently, the research team can time-sample 4–10 non-case controls with EMA assay results concurrent with each case's hazard interval. This between-subject contrast would provide case–control evidence for or against causal inference. There is no need to harvest data from all 1,000 EMA records as might be gathered in a large prevention trial of this type. As mentioned earlier in this section, relative statistical power and precision for the study estimates can be optimized with as few as four non-case controls per newly incident case, and there certainly would be no need for more than 8–10 non-case controls per newly incident case in each time-sampled risk set.

For prevention scientists interested in the epidemiologic nested case–control design with incidence density sampling, the Gordis (2009) and Szklo and Nieto (2007) provide an introduction and intermediate-level coverage, respectively. A basic introduction to the case–control approach as applied in the context of neuropsychiatric research exists (Anthony, 1988). A recent and informative example pertinent to the prevention sciences has been contributed by Weinstein et al. (2011), who studied the possible risk-reducing effects of circulating 25-hydroxyvitamin D by nesting a case–control study within their ongoing Alpha-Tocopherol Beta-Carotene Cancer Prevention Study of Finnish male smokers. Controls were alive and cancer-free at the follow-up time point of diagnosis for each newly incident case. For the case and non-case control in each risk set, the vitamin D assay was performed on a blood specimen that had been drawn and stored at the time of baseline recruitment and assessment in the trial.

Astute readers may appreciate how epidemiologic case–control research with incidence density sampling can be conducted to bring into play variables that were not contemplated at the origination of the prevention research project, as illustrated in the example of an autism-associated locus on chromosome 2. The same approach can be used to measure suspected influences that are too expensive or logistically infeasible to measure for all participants at the time of baseline sampling,

recruitment, or assessment in prevention research but that are deemed to be important to study, as illustrated in the vitamin D example. Of course, the variables measured in this incidence density sampling approach must not be "reactive" in that their values should not change in response to the pathological process or disease state transition under study. The possibility of a reactive covariate is not unique to the case–control design, however; reactive covariates can surface even in properly conducted prospective and longitudinal or developmental research (e.g., as when a child's behavior evokes a parenting response, in a challenge to the idea that the child's behavior always is a consequence of the parenting behavior). The incidence density sampling of non-case controls at the time of discovery of each newly incident case does not finesse that issue of temporal sequencing or feedback looping of this type, when it is present.

An additional note might be of use. Namely, it is allowable for a non-case control to be sampled as a non-case control at one point in time and then to become a newly incident case at some future point in time (i.e., the case–control design with incidence density sampling allows the same participant to be counted as a non-case control at time t and to be counted again as a newly incident case at time $t + k$, where k is an integer from 1 onward). Appropriate statistical approaches for this type of situation are introduced by Szklo and Nieto (2007).

In sum, the epidemiologic case–control design with incidence density sampling can be of special value in prevention research when new research, after baseline, discloses important determinants of the trial outcome that were not well measured at the time of sampling, recruitment, and initial assessment or that were too expensive or logistically infeasible to complete at that time. In many instances, there is no need to measure 100 % of the follow-up participants. Instead, the newly incident cases and a time-matched sample of non-case controls can be assessed, where time matching refers to the situation in which each newly incident case at time t is matched to individuals who have not become newly incident cases as of that time point. When this approach is used, the non-case control sampled at one point in time is allowed to become a newly incident case at a later point in time, and the "dual purpose" appearance of controls who become cases can be accommodated during the steps of statistical analysis and estimation of the cause–effect relationships.

Classical Case–Control Research Design

The epidemiologic case–control research design in its classical form is a between-subject design that does not crunch the time dimension to its zero limit, as is done in the case-crossover research (with focus on the "hazard" interval just before the acute event) and as is done in the nested case–control study with incidence density sampling (with focus on the sampling of non-case controls at the moment in time when each newly incident case is detected). Instead, the classical case–control design involves waiting until a specified time point has been reached and sufficient

numbers of newly incident cases have been detected, at which point there is a contrast of all detected cases at the designated time *t* versus all non-case controls who survived to time *t* without having become cases.

In general, the incidence density sampling approach is preferred, but in some circumstances the classical form is required. To illustrate, consider our research group's hypothesis that sustained use of aspirin and other nonsteroidal anti-inflammatory drugs might delay onset or prevent late-life dementia of the Alzheimer type (DAT), which originated with preclinical research and an initial cross-sectional case–control study of classical form that we had conducted in Australia during the late 1980s (Anthony et al., 2000; Henderson et al., 1992; Zandi et al., 2002; Zandi, Breitner, & Anthony, 2002). One threat to the validity of the research in these examples involves diagnostic assessment of DAT, which optimally includes postmortem examination of brain for detection of the Alzheimer-type lesions in the form of neuritic plaques and neurofibrillary tangles, with concurrent investigation of cerebral infarcts and other vascular lesions (e.g., see Mortimer, 2012). At present, the only viable large sample option in this line of research involves following the participants until death and postmortem examination to combine clinical and pathological findings, with no possibility of time sampling except on the basis of the initial appearance of clinical features.

Another circumstance that might motivate use of the classical case–control study design in prevention research involves the occurrence of unexpected adverse events that might be attributed to the experimental intervention. Early during the post-baseline follow-up interval, these unexpected events might arise sporadically, with no clear connection to the experimental condition until more time has passed. The discovery of the events might be post hoc, perhaps during quarterly, semiannual, or annual meetings of the trial's Data Safety Monitoring Board (DSMB), and might lead to termination of the trial once the balance of evidence indicates more harm than good. This apparently was the situation in a recently terminated prevention experiment sponsored by the National Institute of Neurological Diseases and Stroke, known as the Stenting and Aggressive Medical Management for Preventing Recurrent Stroke in Intracranial Stenosis (Leung, Yu, & Wong, 2011). Interim and final analyses conducted by DSMB had the character of time-to-event analyses, but the discovery of what might be accounting for the adverse outcomes in the Wingspan Stent arm of the trial (e.g., ruptured blood vessels) would require case–control contrasts of the classical form, conducted at trial termination, among the Wingspan Stent-exposed participants.

It is likely that many prevention researchers already are familiar with the classical form of the epidemiologic case–control study. Here again, Gordis (2009) and Szklo and Nieto (2007) provide basic and intermediate-level coverage. Applications and a discussion of case–control issues in neuropsychiatric and prevention research can be found in more specialized reviews (e.g., see Anthony, 1988, 1990).

The Case–Cohort Variant of the Case–Control Research Design

Developmental studies and prevention trials generally start with a very intensive assessment, or with specimen collection of a special character, intended to serve the purpose of a preplanned contrast between subgroups, to investigate possible subgroup variation in effect estimates, or to bring under control a suspected confounding variable that otherwise might be left uncontrolled. When these study procedures are inexpensive or can be completed with ease, the research team typically harvests the information to make the preplanned contrasts for the entire sample under study. Nonetheless, on occasion, the study procedures are prohibitively expensive, too expensive to be completed for 100 % of the participants. Alternatively, completion of 100 % of the assays may be especially time consuming, perhaps more time consuming than the period of funded research permits; or, in the research team's forecast of the occurrence of target outcomes, it becomes apparent that the events will occur with low frequency early in the post-baseline follow-up and with increasing frequency as time passes. This is a typical pattern for drug prevention research that starts in late childhood, when newly incident drug use is rare, and continues to mid-adolescence, when newly incident drug use becomes more frequent (e.g., see Kellam & Anthony, 1998). Under these circumstances, there may be too little time to plan for incidence density sampling of non-case controls at the time each newly incident case is detected. For example, the throughput for a laboratory assay (e.g., for a genetic susceptibility trait) might be too slow to manage the flow of newly detected cases and non-case controls in the final years of follow-up when incidence rates often have accelerated for strongly age-dependent outcomes (e.g., tobacco smoking incidence in mid-adolescence just before age 18, the minimum legal tobacco purchase age in the United States, incidence of Alzheimer-type dementia between ages 80 and 90).

Under these circumstances, the case–cohort variant of the case–control design may be useful. In this approach, there is a random or probability subsampling of all participants at the start of the study, and the laboratory assay throughput begins with pre-scheduling work on this subsample, well before any newly incident cases have been detected so that the assay workflow can be distributed throughout the entire study interval. Some subsequently detected cases are members of the probability subsample, and their specimens will be processed part and parcel with the pre-scheduled workflow. As for newly incident cases not included in the initial probability subsample, their specimens are added to the workflow as soon as the newly incident case is detected, most often with the laboratory personnel being masked or blinded to case and non-case status. The distribution of the subsample workflow throughout the entire study interval guarantees that there will be a mixture of non-case specimens and case specimens even during the final months of the laboratory work. (Otherwise, without balanced distribution of the subsample work, the only specimens processed during the final months would be specimens from newly detected cases, with potential differential measurement bias.)

The Centers for Disease Control and Prevention recently launched its Pregnancy and Influenza Project with a case–cohort design to probe the effectiveness of maternal influenza vaccination and real-time reverse transcription polymerase chain reaction assays to confirm influenza infections of pregnant women and their newborn infants (Thompson et al., 2011). Readers interested in an illustration of a case–cohort study nested within a long-term developmental research plan may wish to refer to Agudo and colleagues (2012), who recently completed a case–cohort study of aromatic DNA adducts pertinent to prevention of gastrointestinal cancers. Li and colleagues (2012) are responsible for a recent overview of the statistical approaches used in case–cohort research of this type, including a description of a statistical innovation for analyses of case–cohort data that yokes the piecewise exponential approach with Poisson regression and Taylor series linearization. The data analyzed by Li and colleagues are from the Finnish Alpha-Tocopherol Beta-Carotene Cancer Prevention Study mentioned in Sect. 17.3.

This chapter's description of the case–cohort approach may lead some readers to think that the approach is valuable only when there are intensive "wet bench" laboratory assays in the list of study procedures. To illustrate a hypothetical example with a more social psychological or social science character, consider a "mixed method" prevention trial that begins with recruitment of participants who are willing to share their Facebook timelines with the research team so that intensive qualitative assays might be performed using Ethnograph® or other similar software for analysis of individual characteristics and between-individual interaction patterns or processes that might be predictive of the trial outcomes or that might modify the effects of the preventive intervention (e.g., see Gajaria, Yeung, Goodale, & Charach, 2011, for a related observational study approach). Alternatively, there might be a videotape of a structured transactional task session with parents and children interacting within an extended baseline session, intended to yield enriched baseline data on transactions of importance in an evaluation of trial effects. The throughput time for rigorous independent coding of these videos by pairs or sets of raters tends to be lengthy, as we have found in our observational studies of mothers and children (e.g., see Johnson et al., 2002; Mansoor et al., 2012). Large sample versions of prevention trials with baseline assessments that require intensive standardized coding of qualitative material of this type may well benefit from an application of the case–cohort variant of the case–control approach (e.g., see Wadsby, Sydsjo, & Svedin, 2001, for a small sample intervention trial of this type).

Conclusions

In sum, it may be said that the field of prevention science has early origins in public health research conducted by epidemiologists and that there are many intersections between the work of public health scientists and the clinical or developmental psychologists who are drawn to the prevention arena (e.g., see Anthony, 1990;

Ialongo 2002). Nonetheless, many prevention scientists have no more than a passing acquaintance with standard epidemiologic research approaches and do not apprehend the applicability of these approaches in the design and conduct of prevention research.

The largest impediment in this regard may involve the debate about discrete categorical and dimensional outcomes in prevention research, with public health and the biomedical sciences tending to be quite comfortable with "diagnostic" decision-making of the type the surgeon must make when deciding to cut or not to cut or prescribing decisions when the medicine might have potentially toxic effects as well as anticipated beneficial effects. This comfort with binary prevention outcomes or other discrete categorical outcomes fits well with ideas about using the case-crossover hybrid design or the variants of the between-subject case–control approach.

Part of the argument of this chapter is that dimensional outcomes, and the processes that lead toward between-group differences in means, also can be studied with an eye that seeks out acute state transitions during those processes. The example of a child's first chance to try a drug has been mentioned as a state transition that can be studied even when the prevention trial focuses on drug involvement as measured in a dimensional fashion. Inspecting these processes, it often is possible to spot the state transitions. Once a state transition has been discovered in a developmental process, the epidemiologic case-crossover and case–control research approaches become viable tools for prevention research.

As stated in Sect. 17.1, if the evidence from prevention research is to be definitive, compelling, and translational into public health action, our studies must be tightly controlled and disciplined, perhaps even rigid. Even so, innovation in research approach is required to secure funding to complete prevention research with lasting impact. One hopes that the ideas presented in this chapter may lead future prevention scientists to design and conduct their future prevention trials with innovative features of the epidemiologic case-crossover and case–control designs described in this chapter.

Acknowledgments The preparation of this chapter was supported, in part, by a National Institute on Drug Abuse (NIDA) senior scientist and mentorship career award (K05DA015799). There are no other conflicts of interest to disclose. The content is the sole responsibility of the author and does not necessarily represent the official views of Michigan State University, NIDA, or the National Institutes of Health.

References

Agudo, A., Peluso, M., Munnia, A., Luján-Barroso, L., Sánchez, M. J., Molina-Montes, E., et al. (2012). Aromatic DNA adducts and risk of gastrointestinal cancers: A case-cohort study within the EPIC-Spain. *Cancer Epidemiological Biomarkers and Prevention, 21*(4), 685–692.

Anthony, J. C. (1988). The epidemiologic case–control strategy, with applications in psychiatric research. In A. S. Henderson & G. D. Burrows (Eds.), *Handbook of social psychiatry* (pp. 157–171). Amsterdam: Elsevier Science Publishers.

Anthony, J. C. (1990). Prevention research in the context of epidemiology, with a discussion of public health models. In P. Muehrer (Ed.), *Conceptual research models for preventing mental disorders (pp. 1–32) (DHHS Publication No. ADM 90-1713)*. Washington, DC: GPO.

Anthony, J. C., Breitner, J. C., Zandi, P. P., Meyer, M. R., Jurasova, I., Norton, M. C., et al. (2000). Reduced prevalence of AD in users of NSAIDs and H2 receptor antagonists: The Cache County study. *Neurology, 54*(11), 2066–2071.

Dodge, K. A., Malone, P. S., Lansford, J. E., Miller, S., Pettit, G. S., & Bates, J. E. (2009). A dynamic cascade model of the development of substance use onset. *Monographs of the Society for Research on Child Development, 74*(3), vii–119.

Gajaria, A., Yeung, E., Goodale, T., & Charach, A. (2011). Beliefs that attention-deficit/hyperactivity disorder and response to stereotype: Youth posting in Facebook groups. *Journal of Adolescent Health, 19*(1), 15–20.

Gordis, L. (2009). *Epidemiology* (4th ed.). Philadelphia: W. B. Saunders.

Henderson, A. S., Jorm, A. F., Korten, A. E., Creasey, H., McCusker, E., Broe, G. A., et al. (1992). Environmental risk factors for Alzheimer's disease: Their relationship to age of onset and to familial or sporadic types. *Psychological Medicine, 22*(2), 429–436.

Ialongo, N. (2002). Wedding the public health and clinical psychological perspective as a prevention scientist. *Prevention and Treatment, 5*(1), Article 4.

Johnson, A. L., Morrow, C. E., Accornero, V. H., Xue, L., Anthony, J. C., & Bandstra, E. S. (2002). Maternal cocaine use: Estimated effects on mother-child play interactions in the preschool period. *Journal of Development and Behavioral Pediatrics, 23*(4), 191–202.

Kellam, S. G., & Anthony, J. C. (1998). Targeting early antecedents to prevent tobacco smoking: Findings from an epidemiologically based randomized field trial. *American Journal of Public Health, 88*(10), 1490–1495.

Leung, T. W., Yu, S. C., & Wong, K. S. (2011). Have medical therapy and stenting been fairly compared? A repercussion upon termination of recruitment in the SAMMPRIS trial. *International Journal of Stroke, 6*(4), 312–314.

Li, Y., Gail, M. H., Preston, D. L., Graubard, B. I., & Lubin, J. H. (2012). Piecewise exponential survival times and analysis of case-cohort data. *Statistics in Medicine, 31*(13), 1361–1368.

Maclure, M. (1990). The case-crossover design: A method for studying transient effects of brief exposures on the risk of rare acute events. *American Journal of Epidemiology, 132*(4), 781–782.

Mansoor, E., Morrow, C. E., Accornero, V. H., Xue, L., Johnson, A. L., Anthony, J. C., et al. (2012). Longitudinal effects of prenatal cocaine use on mother-child interactions at ages 3 and 5 years. *Journal of Developmental and Behavioral Pediatrics, 33*(1), 32–41.

Mortimer, J. A. (2012). *The Nun Study: Risk factor for pathology and clinical-pathologic correlations*. UK: Current Alzheimer Research. PMID: 22471869.

O'Brien, M. S., Comment, L. A., Liang, K. Y., & Anthony, J. C. (2012). Does cannabis onset trigger cocaine onset? A case-crossover approach. *International Journal of Methods in Psychiatric Research, 21*(1), 66–75.

Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine, 336*(7), 453–458.

Sanders, S. J., Murtha, M. T., Gupta, A. R., Murdoch, J. D., Raubeson, M. J., Willsey, A. J., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature, 485*, 237–241. doi:10.1038/nature10945.

Szklo, M., & Nieto, J. (2007). *Epidemiology: Beyond the basics*. Sudbury, MA: Jones & Bartlett.

Thompson, M., Williams, J., Naleway, A., Li, D. K., Chu, S., Bozeman, S., et al. (2011). The pregnancy and influenza project: Design of an observational case-cohort study to evaluate influenza burden and vaccine effectiveness among pregnant women and their infants. *American Journal of Obstetrics and Gynecology, 204*(6 Suppl 1), S69–S76.

Wadsby, M., Sydsjo, G., & Svedin, C. G. (2001). Evaluation of an intervention programme to support mothers and babies at psychosocial risk: Assessment of mother/child interaction and mother's perception of benefit. *Health and Social Care in the Community, 9*(3), 125–133.

Warner, L., Macaluso, M., Austin, H. D., Kleinbaum, D. K., Artz, L., Fleenor, M. E., et al. (2005). Application of the case-crossover design to reduce unmeasured confounding in studies of condom effectiveness. *American Journal of Epidemiology, 161*(8), 765–773.

Weinstein, S. J., Yu, K., Horst, R. L., Parisi, D., Virtamo, J., & Albanes, D. (2011). Serum 25-hydroxyvitamin D and risk of lung cancer in male smokers: A nested case–control study. *PLoS ONE, 6*(6), e20796. doi:10.1371/journal.pone.0020796.

Wu, L. T., & Anthony, J. C. (2000). The use of the case-crossover design in studying illicit drug use. *Substance Use & Misuse, 35*(6–8), 1035–1050.

Zandi, P. P., Anthony, J. C., Hayden, K. M., Mehta, K., Mayer, L., Breitner, J. C., et al. (2002). Reduced incidence of AD with NSAID but not H2 receptor antagonists: The Cache County Study. *Neurology, 59*(6), 880–886.

Zandi, P. P., Breitner, J. C., & Anthony, J. C. (2002). Is pharmacological prevention of Alzheimer's a realistic goal? *Expert Opinion on Pharmacotherapy, 3*(4), 365–380.

Chapter 18
The Use of Simulation Models in Public Health with Applications to Substance Abuse and Obesity Problems

David Levy

Introduction

Evaluation studies often focus on a particular policy applied to a specific population. Due to data limitations, purely statistical methods of analysis are often ill equipped to simultaneously consider all of the important relationships inherent in complex problems, such as the effects of tobacco control policies on smoking rates at different ages and how those effects impact health outcomes over time. An alternative, complementary approach to purely statistical modeling is simulation modeling (SM), also known as computational models.

Rather than a piecemeal approach focusing on one aspect of the problem, SMs attempt to provide a more comprehensive, whole-system perspective. A conceptual model of the problem is generally developed before the data are collected to structure the problem, to clarify data needs, and to identify gaps in our knowledge base. Modeling methods are designed to account for complexities with their capacity to examine and identify nonlinear patterns, time-delayed effects, multiple interacting factors, and feedback loops.

Computer simulation is now widely used to understand and predict behavior in complex systems, such as financial markets, ecological and biological systems, and traffic patterns. The use of SMs in public health is at a nascent stage but has tremendous potential to bridge the gap between research and practice (Homer & Hirsch 2006; Levy, Bauer, & Lee, 2006; Sterman, 2006). SMs combine information from different sources to provide a useful tool for examining how the effects of public health policies unfold over time in complex systems and impact population health (Homer & Hirsch, 2006; Levy et al., 2006).

D. Levy (✉)
Department of Oncology, Georgetown University, 3300 Whitehaven Street, NW Suite 4100, Washington, DC 20007, USA
e-mail: dl777@georgetown.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_18,

In the preventions area, models focus on the role of risk factors, such poor diet, lack of physical activity, alcohol use, smoking, or illicit drug use. For example, SMs of smoking behaviors date back at least to the 1980s (Roberts, Homer, Kasabian, & Varrell, 1982). More recently, models of the effect of tobacco control policies, trends in smoking prevalence, and health outcomes have been developed by Mendez and colleagues (Mendez & Warner, 2004; Mendez, Warner, & Courant, 1998), Orme and colleagues (Orme et al., 2001), Tengs and colleagues (Tengs, Ahmad, Moore, & Gage, 2004; Tengs, Osgood, & Lin, 2001), and Levy and colleagues (Levy, Chaloupka, Gitchell, Mendez, & Warner, 2002; Levy, Cummings, & Hyland, 2000a, 2000b). The *SimSmoke* model of Levy and colleagues considers a broader array of public policies than the other models (Levy et al., 2000a, 2000b; Levy, Chaloupka et al., 2002) and will be used extensively in examples throughout this chapter along with work in a relatively new but important area—obesity.

This chapter begins with a discussion of the uses of simulation models, then discusses different types of simulation models, followed by the steps involved in developing different models. The paper then switches gears by summarizing some of the simulation literature and identifying salient issues in terms of the stages involved from policy to risk behaviors to health outcomes.

Simulation Models: Background

Aim of Simulation Models

SMs rely on mathematical equations to map the quantitative relationships between a series of inputs and one or more outputs, such as where smoking is an input and various health outcomes, such as heart disease and lung cancer, are the outputs. Outputs from one set of equations can be used as inputs to a second set of equations. For example, smoking rates might be explained by the social environment and/or environmental changes, then smoking as an output can be used as an input to determine health outcomes. Equations linked in this way form the basis of SMs.

As shown in Fig. 18.1, the effects of public policies can be seen as involving stages; public policies affect environments and knowledge, which affect adverse behaviors, which in turn influence health outcomes. For example, a policy aimed at youth purchase of cigarettes might affect the retailing environment and the social context of the acceptability of smoking, which in turn affects smoking initiation. Smoking, if maintained over time, ultimately affects the likelihood of adverse health outcomes at later ages (e.g., through disability or death due to smoking-related diseases). In the case of obesity-related behaviors, policies can affect the eating or physical activity environment, which in turn affects nutrition and physical activity. The effect of these behaviors on health outcomes can occur through body mass index (BMI) or through consumption of unhealthy foods (e.g., consumption of

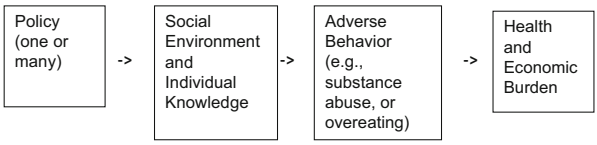

Fig. 18.1 The stages from public policy to public health outcomes

sugar-sweetened beverages (SSBs) may directly affect the likelihood of diabetes). Models can encompass each of these stages or a subset of stages. At each stage, the effects may depend on moderating factors, such as the age, gender, racial-ethnic identity, or socioeconomic status of the individual.

To capture the "big picture," SMs generally rely on data from a variety of disciplines and sources. Rather than attempt to estimate all the pathways in a single statistical equation or group of equations, SMs systematically combine information from different studies with plausible theoretical relationships. Model building is a cross-disciplinary venture that combines the theory and evidence from diverse fields, such as statistics, epidemiology, biology, nutrition, sociology, psychology, economics, and programming (Huang & Glass, 2008). The conceptualization or framework for different stages of the model may be based on different disciplines (e.g., economics or sociology can inform the policy analysis, biology can inform the role of genetics, and epidemiology can inform the patterns of risk factors or diseases over time).

Once developed, the model is then validated against actual trends and turning points in trends for the population as a whole and for various population subgroups. As new empirical studies are conducted, information from those studies can be incorporated in the model, and the validation exercise begins again.

Uses of Simulation Models

While most statistical studies rely on one specific data set applied to a specific problem, SMs often bring together information from different data sources and from different, but related, studies. Models can be used to explain discrepancies in different studies, by considering the underlying differences in those studies in a broader framework. Indeed, SMs can be used to synthesize knowledge from a variety of disciplines and systematically combine that information. Thereby, SMs often reveal limitations in our knowledge and suggest important areas for study. Furthermore, SMs can help us understand the complexity of behavior in a larger system rather than pieces of the system in isolation and, thereby, provide a heuristic method for understanding public health systems and providing solutions.

At an applied level, SMs can be used to teach and understand. While there has been an increasing interest in incorporating empiricism into policy-based decisions, policy makers may lack the necessary time, expertise, and financial resources to obtain and fully understand the numerous articles available in academic journals. SMs combine the best known information and, thereby, serve as a summary of

knowledge. SMs may also be considered at a more academic level of a means of extending our knowledge and as a consequence help us understand limitations of our knowledge. By explicitly considering the links between different stages from policies to risk behaviors to health outcomes, the links are made explicit. By considering these interrelationships over time, the manner in which of the stages unfold over time can be better understood.

At least two other uses can be identified for smoking SMs (Levy et al., 2002). First, they demonstrate the need for public health policy by showing likely future trends in smoking and the nature and extent of the effects of smoking on health and other outcomes over time. For example, the original intent for developing *SimSmoke* was to provide justification for tobacco control to policy makers. The *SimSmoke* model shows future trends in smoking, the associated health outcomes, and the effects of public policies on those outcomes. By projecting the number of lives saved from implementing polices, tobacco control activities can be better justified. The model can also be used for planning purposes (how to refine existing policies, the need for additional policies, or particular demographic groups to target).

Where sufficient data are available, the model tracks the effect of policies from a baseline year before major policy changes are implemented. The *SimSmoke* model has been applied to the United States as a whole, single states within the United States (i.e., Arizona, California, and New York), and other sovereign nations. Predictions of the smoking rate are tracked from the baseline year to the most recent year for which data are available and preferably a period before major changes in public policy occurred. Smoking rates change due to the smoking patterns of different cohorts of smokers and due to the effects of the different policies implemented during the tracking period. Once the model is developed for a particular nation or state, trends in smoking rates in the absence of policy change can be predicted by maintaining policies at their baseline level over the tracking period. This provides the counterfactual of what would have happened in the absence of policies and, thus, by comparing actual rates with policies actually implemented can be used to gauge the effect of past policies since the tracking period began. By then programming the actual changes in individual policies, the effect of policies individually and in combination can be predicted (Levy et al., 2010b). The success and failure of past policies individually and collectively can then be shown.

Another use of SMs is forecasting and planning. They can be used to forecast health needs arising from obesity. By using SMs to estimate the effects of different policies individually and in various combinations, an effective strategy can be developed. SMs can also show the effect of policies on those outcomes through "what if" scenarios that model various policies under specific assumptions—how much smoking rates would be reduced if tobacco control policies were implemented individually with other policies held constant or with policies simultaneously implemented in some combination? SMs may be used to identify "leverage points" in the system—where the most change in the system (in the desired direction) can result from the least amount of resource or effort applied.

Since random assignment is usually not possible, SMs are most commonly applied in prevention science to population-level policies and laws. They may also be applied to universal preventive interventions that are commonly informed by the developmental epidemiology of the development of risk and protective factor (and thus may target early proximal targets such as aggressive behavior in first grade) or later proximal targets in middle school and high school focusing on normative beliefs and expectancies. For each type of intervention, the ultimate public health outcomes, especially those relevant to cost-effectiveness analysis such as health and economic outcomes, often occur many years after a study is completed. Modeling may be used to follow a cohort of individuals to estimate the effect of interventions on these outcomes and to generalize the results of a particular random control study to a broader population. Similarly, for community prevention that focuses on contextual aspects in the school and community context, modeling may be used to estimate outcomes occurring after a study has been completed and to generalize the results to other populations.

Ideally, the model can become a central part of a surveillance and evaluation system. Models depend on the availability of appropriate data. Where data have not been systematically collected, the model suggests the type of data that is needed (e.g., the specific questions asked in surveys) to improve surveillance systems for monitoring, such as the appropriate measure of smoking rates, the amount of detail by sociodemographic status, and the relevant aspects of a fully implemented policy. For example, models can help determine appropriate goals (e.g., *Healthy People 2020*) and likely progress toward health objectives (Levy et al., 2010a; Levy, Nikolayev, Mumford, & Compton, 2005; Levy, Tworek, Hahn, & Davis, 2008). Thus, SMs can provide guidance on the data needed, not only for the model, but for formulating policy. Models may also provide guidance for evaluating policies. The model may help distinguish the factors relevant to measuring policies (e.g., the characteristics relevant to policy success and the role of other policies), the outcomes to be considered (e.g., the specific populations likely to be affected and time pattern of effects), and whether the results are plausible in light of past studies.

Like the model itself, the informational process involves adjustments in a dynamic context. As new surveillance data are collected, the improved measures (e.g., the specific measures of smoking uptake) are inputted for better accuracy. As new evaluation studies are conducted, these studies may be used to revise policy effect size parameters in the model. With these improvements, the model provides a sounder basis for planning.

Types of Simulation Models

SMs can be static and consider snapshots usually at two points in time or two different scenarios at a point in time, or they can be dynamic and consider the path of changes over time. Since most models are dynamic, the discussion here focuses on dynamic models.

Dynamic models may be further classified by other dimensions: whether they are macro or micro models, whether they simulate events among a single cohort of individuals over their lifetime or among the entire population over a specified period, whether events are deterministic or stochastic, whether events occur in continuous or discrete time, and whether they are unidirectional or system dynamics.

In general, the primary difference between macro models or compartmental models as they are now most often known and individual models is the degree of complexity built into the model. By definition, macro (aka compartmental) models aggregate individuals into compartments (i.e., groups) and thus do not track individual differences. Groups may be defined by sociodemographic status—race, gender, age, education, and income (e.g., a group for African-American men aged 60–70). Sociodemographic groups may be further distinguished by smoking status (e.g., whether a person smokes), the extent of their smoking (every day or some days or by quantity smoked), the duration of their smoking (in terms of either age of initiation or actual years smoked), or type of cigarette smoked (menthol vs. -non-menthol, other cigarette characteristic, or brand). In models with health outcomes, former smokers are distinguished from current and never smokers because, while their health risks are lower than those for smokers, they are higher than for those never smokers. Former smokers may be further classified by years quit and quantity and duration of smoking before quitting, since these factors have been found to affect health outcomes.

In contrast to macro models, *individuals* (rather than groups) are modeled in micro models. Microsimulation modeling techniques were first proposed as a method for examining the effects of policy change in the late 1950s by Orcutt (1957). Orcutt envisioned microsimulation modeling as a broadly applicable technique, with models simulating the effects of policy changes on simulated decision-making units that could include individuals, families, or corporations. During the 1970s microsimulation models were developed to guide US social policy decisions, for example, models of traffic flow are used to plan transit projects (Lemp, McWethy, & Kockelman, 2007). Models of the transmission of infectious diseases such as HIV (Cassels, Clark, & Morris, 2008) and influenza (Ackerman, Longini, Seaholms, & Hedin, 1990) were used to examine the effects of intervention and policy change on disease transmission.

In micro models, the transitions of agents are often structured by rules of behavior. Agent-based models (Bonabeau, 2002) employ rules for individual (or "agent") behavior and interactions with other agents. For example, economic models focus on individual behavior and usually apply maximizing behavior subject to income or wealth constraints (e.g., to show the effects on obesity of price). Sociological and infectious disease models tend to focus more on social interaction. For example, interactions may be based on location, as in network models, where behaviors are transmitted across individuals who are in close proximity in terms of either geographic location or social interaction. These are similar to artificial neural networks, a type of computational model inspired by biological neural networks, in which there is an interconnected group of artificial

neurons that uses a connectionist approach to computation. While models tend to apply simple models of behavior (the KISS concept of "Keep It Simple, Stupid"), they may allow for adaptive behavior over time, similar to an adaptive artificial neural system, which changes its structure based on external or internal information that flows through the network during the learning phase.

In micro models, individuals are generally mapped from a multivariate distribution of sociodemographic and behavioral characteristics. For example, in a smoking model, a distribution of gender, socioeconomic status might be used to distinguish smoking patterns, or in a model of obesity, a distribution of BMI may be applied to different age, gender, and racial-ethnic groups. Thus, micro models enable a finer distinction on population characteristics but also require more knowledge about those characteristics and the interrelationship between characteristics. One frequently used type of probabilistic method is the Monte Carlo Model, which draws randomly from a fixed distribution of population characteristics.

When models involve multiple equations, the outputs of one equation may affect the inputs for another equation, whose outputs, in turn, affect inputs of the first equation. These feedback loops are one hallmark of systems dynamic models. This approach was developed by Forrester (Homer & Hirsch, 2006). The central tenet is that complex behaviors of organizational and social systems are the result of ongoing accumulations (e.g., of people of various classifications) both balancing and reinforcing feedback mechanisms. For example, policies may affect smoking rates, which in turn may affect social norms or the ability to implement policies. In the absence of feedback loops, the models are unidirectional or deterministic. Because feedback loops add considerably to the complexity of the model, systems dynamic models are often developed as macro as opposed to micro models.

Steps in the Development of a Model

Computational (or mathematical) modeling is at its core a method to simplify reality and create a more tractable version of it, thereby making the problem more amenable to inspection. The simplified problem can be probed in the virtual environment using the scientific method and has the advantage of doing it faster, cheaper, and without the harms/risks associated with the real world. The most useful models are those that retain only the most salient aspects of reality, reducing relationships to mathematical equations that reflect the essence of the modeler's understanding of the problem. The model reflects the best available evidence as well as a variety of assumptions. Once the model is built and validated, it can be used to generate estimates of how pertinent outcomes of interest will unfold over time given the parameters and initial starting values ascribed to the model by its builder. This process of simulation (usually with the aid of a computer) shares much in common with its laboratory counterparts (i.e., any number of variables can be adjusted in a controlled fashion and the resulting outcomes observed).

The development of an SM generally involves several steps: (1) define the problem and its scope; (2) determine the most appropriate type of model and basic structure; (3) specify the equations of the model; (4) validate the model; (5) record output; and (6) conduct sensitivity analyses.

A basic structure is determined in choosing between micro and macro modeling and the number of stages that are upstream (e.g., from BMI back to diet to policy) and downstream (e.g., from weight to health outcomes). The choice is dependent on the problem being considered, the data available, and the tension between the simplicity and complexity of the model. A more complex model can better represent reality, but simpler models are easier to describe and thus more transparent and include fewer parameters that need to be informed.

Three steps are essential in conceptualizing the equations: (1) identifying a fixed number of distinct states and characteristics associated with these states, which requires determining the level of heterogeneity in the modeled population; (2) specifying rules for transition through those states; and (3) setting values of model parameters that link inputs and outcomes. Initial model parameters may be statistically estimated or derived from other, usually published, sources. It is not usual to have gaps in the information used to quantify relationships. Best practice is to use the strongest evidence available, but when solid empirical evidence is not available, expert opinion can be used to estimate the parameters of the model (Swinburn, Gill, & Kamanyika, 2005). Model parameters may be selected so that the model can reproduce expected or observed results, a process called "calibration," or they may flow directly from available data. Distributions can be specified for those parameters for which the evidence is less certain.

Both macro and micro models require assumptions about the relevant sociodemographic and smoking characteristics to consider. The choice of characteristics considered depends on three considerations. First, it is useful to consider the potentially relevant characteristics, which are determined by the issues thought to be of key importance to the problem at hand. For example, what are the relevant characteristics that affect the link between behaviors and health outcomes—age, gender, and socioeconomic status? Which sociodemographic characteristics are associated with different smoking patterns? Do the effects of public policies vary by sociodemographic and smoking characteristics? Second, the choice of characteristics incorporated into the model depends on data availability (i.e., can sufficiently accurate data be obtained that distinguish the population by the potentially relevant characteristics). Finally, the choice of characteristics considered will also affect the degree of complexity of the model. Again, a trade-off is usually encountered between model simplicity (considering fewer characteristics) and incorporating potential complexities. A macro model may be more complex in terms of the number of difference classifications of groups considered, while the micro model may be more complex in terms of its populating the model by individuals rather than groups.

In dynamic models, populations are usually distinguished at minimum by age to represent the relevant transitions over time, such as progressing from age 10 in the year 2010 to age 11 in 2011. Discrete time models allow individuals to move

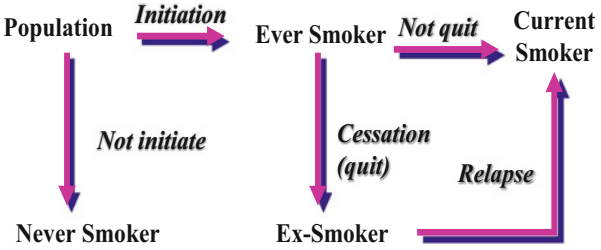

Fig. 18.2 Smoking model: evolution of smokers

between states (e.g., from age a in period t to age $a + 1$ in period $t + 1$) only at fixed time intervals (possibly dependent on the state), with transition probabilities specific to the cycle length (e.g., month or year). In continuous models, intervals between transitions have continuous distributions. Most dynamic models are set up as discrete models by year, corresponding to availability of data which is often annual.

Over time, the transitions depend on transitions rules. In micro models, transition rules may apply to specific behaviors of individuals, whereas in compartmental models the transitions focus on transitions by age (through births and deaths), by smoking status (from never smokers to current smokers and then possibly to former smokers and possibly to relapse), by how the effects policies unfold over time (depending on how they are implemented and the other policies in effect), and by health outcomes (e.g., by degrees of morbidity and death). Transition rules may depend on individual or group characteristics in either type of model. However, to reduce complexity, they are usually specified as Markov (Cox & Miller, 1965; Sonnenberg & Beck, 1993) whereby they depend only on the current state (e.g., by carrying all past information forward into the current state).

Central to smoking models is the transitions into and out of smoking, as shown in Fig. 18.2. Transitions into smoking are defined by initiation rates, usually defined in terms of the percentage of never smokers in a particular period t, which become smokers by period $t + 1$. The definition should correspond to the definition of smoker tracked at later ages. For example, if the model considers health effects in terms of established smokers, defined in terms of those who have smoked 100 cigarettes in the their lifetime and are currently smoking, then initiation rates must ultimately explain the transition to established smoker. Smoking models also consider cessation from smoking and often relapse back into smoking. Cessation is generally considered over a fixed period such as the last year, whereas relapse may depend on the number of years since quitting. Cessation and relapse may be collapsed into a measure of cessation net of relapse for purposes of simplicity, when the issues surrounding relapse are not considered central to the problems being considered by the model.

Validation is the process of assessing whether the model is consistent with data not used for calibration, an independent verification process also called "external validation." By showing that the simulated outcomes of the model conform to historical data, the credibility of the information (e.g., policy evaluation studies)

used in developing the model is enhanced. Once the *SimSmoke* model is developed for a particular nation, levels and trends in smoking prevalence over the tracking period are compared with actual data, preferably from large-scale surveys or studies. Comparisons are made for the population as a whole and for specific demographic groups. Using mean square error and other criteria, the comparisons serve to validate the model for future projections and determine weaknesses of the model. Similarly, predictions can be validated in terms of health outcomes or intermediate effects of policies, such as attitudes.

Sensitivity analysis refers to the presentation of model results under different values of model parameters that are subject to uncertainty or different features in the design of the model (e.g., what if interventions were delivered by a practice nurse rather than a general practitioner?) In particular, policy effect sizes can be considered over the range of estimates found in the better studies. Sensitivity analysis shows the uncertainty over plausible assumptions and parameters.

Review of Current Prevention Models

The ultimate goal, from a public health standpoint, is to reduce the incidence and costs of disease associated with risk factors. By beginning with poor health and other social problems as the ultimate concern, the aspects of risk behaviors that are important to these outcomes can be identified. Thus, we begin by reviewing SMs of how risk factors affect health and related social problems. We then consider models projecting trends in risk factors, beginning with simple models that project future rates of risky behaviors based on past rates, and then consider more complex models that relate risks factors to specific behaviors, societal interactions, or environmental conditions. Finally, we consider the role of policy in affecting those conditions or risk factors.

Health and Related Outcomes

Smoking has been directly linked to heart disease, stroke, lung and other cancers, and emphysema. Similarly, obesity has been associated with a higher risk of heart disease, cancers, and other diseases (Popkin, Kim, Rusev, & Zizza, 2006). Each outcome is associated with different risks of morbidity and mortality as related to the risk factor. The total mortality risks can be developed as a single measure (e.g., by using a single equation that relates smoking or BMI to mortality risk) or by summing over the mortality risks associated with different diseases taking into account their competing effects. When examining mortality, the model may further take into account the average years of life lost at each age (Fontaine, Redden, Wang, Westfall, & Allison, 2003; Peeters et al., 2003). Morbidity varies by the type

and severity of disease. Aggregation over diseases can be in terms of quality-adjusted or disability-adjusted life years lost or in terms of medical costs.

The excess risks of morbidity and mortality due to smoking or obesity are more pronounced at later ages, but health effects arise as early as childhood. The effects may depend on the level of smoking or BMI in a nonlinear fashion (Flegal & Graubard, 2009; Flegal, Graubard, Williamson, & Gail, 2007a, 2007b; Flegal, Williamson, Pamuk, & Rosenberg, 2004) and will vary by age, gender, or socioeconomic status. Smoking risks are related to age of initiation, quantity smoked, and duration of smoking, and the effects may vary by age, gender, race, and socioeconomic status. Smoking has also been indirectly linked through second-hand smoke to heart disease and lung and other cancers. In addition, smoking has been linked to many other diseases, including health outcomes of children whose parents smoke and adverse maternal health outcomes, such as sudden infant death syndrome and low birth weight.

Due to the highly addictive nature of smoking and because smoking generally begins at a young age, smoking patterns remain relatively constant over the lifetime of an individual and are subject to the success of quitting attempts. In examining the effects of changes in dietary and physical activity or the effects of policy on those behaviors, the risks of obesity may be expected to vary more over the life cycle of an individual than those of smoking (Finkelstein et al., 2008). Obesity may affect current and future health, not only because of the correlation of future with current obesity rates as individuals age, but also because obesity may have cumulatively affect risks over time or may have caused a disease in the past that has ramifications for the future (e.g., diabetes). Thus, a life cycle model requires knowledge of how youth BMI translates into adult BMI and the relative health risks associated with different levels of weight at each age as dependent on current and possibly past BMI (Bibbins-Domingo, Coxson, Pletcher, Lightwood, & Goldman, 2007; Finkelstein, Fiebelkorn, & Guijing, 2003; Finkelstein, Trogdon, Cohen, & Dietz, 2009; Lightwood et al., 2009). At some point the increased costs at younger ages may be offset by reduced costs at later ages due to decreased likelihood that an individual will be alive at later ages (van Baal et al., 2008). Given the BMI characteristics and sociodemographic distributions of the population, health outcomes can be related to the distribution of BMI at a point in time (Finkelstein et al., 2003, 2009) but may also depend on past levels of BMI. However, the effect of obesity on these outcomes may also depend on how obesity is measured (Flegal & Graubard, 2009) and how diet and physical activity act independently (e.g., link to mental health) and in combination to determine obesity.

As shown in Fig. 18.3, alcohol use, smoking, illicit drug use, and obesity have also been associated with adverse outcomes other than health, including emotional distress and work productivity (Finkelstein, Ruhm, & Kosa, 2005; Trogdon, Finkelstein, Hylands, Dellea, & Kamal-Bahl, 2008) or, in the case of the youth, educational achievement and psychosocial outcomes. Alcohol and illicit drugs have been associated with an increased propensity for criminal behavior (Miller, Levy, Cohen, & Cox, 2006), although estimates of these propensities are more tenuous

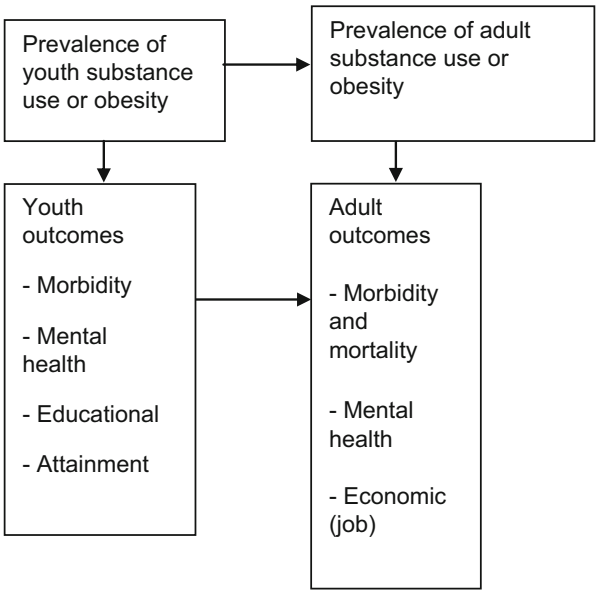

Fig. 18.3 Future public health and economic outcomes

than for health risks. In the case of youth-oriented policies, economic and health effects often occur later in life, when the individual reaches adulthood.

Finally, the effects of a particular risk factor, such as smoking, on health can be considered in a model devoted entirely to smoking or in broader models that simultaneously consider multiple risk factors (Jones et al., 2006; Milstein et al., 2007) and a broad array of diseases (Goldman et al., 2009; Rodgers et al., 2004; Van Meijgaard, Fielding, & Kominski, 2009). In a model of obesity, smoking may cause competing risks or reinforcing effects on mortality (Flegal, 2007; Flegal et al., 2007a, 2007b; Jia & Lubetkin, 2009; Stewart, Cutler, & Rosen, 2009). For Canada, the POHEM Monte Carlo simulation model (Flanagan et al., 2003; Kopec et al., 2009; Wolfson 1994) has been expanded to consider the role of obesity along with other lifestyle factors (e.g., smoking).

Modeling Trends in Risk Factors

For planning purposes, especially to gauge future healthcare needs, simulation models are used to project future trends in health outcomes and how they are influenced by risk factors. Projecting the health effects of risk factors will require knowledge of how risk factors change over time and will require SMs that accurately predict future trends in risk factors.

Dynamic models that project risk factors over time begin with the population distinguished by a minimum by age. Over time, the population evolves through births and deaths and possibly migration into and out of the geographic area.

The model may also distinguish the population at each age by other population characteristics that are likely to affect risk factors over time, such as gender, racial-ethnic group, and socioeconomic status (Beydoun & Wang, 2009; Wang & Beydoun, 2007; Wang, Colditz, & Kuntz, 2007).

More sophisticated models explicitly consider transition probabilities by age (i.e., age i in period to age $i + 1$ in period $t + 1$), where the transitions may be estimated by following age groups over time. As discussed above, smoking models will need to consider initiation, cessation, and relapse rates. Ideally, these can be derived using longitudinal data that are sufficient to distinguish by age and gender, as well as any other important sociodemographic factors. Since sufficient longitudinal data are generally not available, synthetic cohorts are often used (e.g., group by age). For the United States, Basu and colleagues (Huang, Basu, O'Grady, & Capretta, 2009) used Medical Expenditure Panel Survey data (2001–2002 and 2004–2005) to estimate and compare the 1-year transitions across BMI categories for children and adults and validated the model using the 2005–2006 National Health and Nutrition Examination Survey (NHANES).

Data (e.g., due to sample size and period covered) limit the ability to statistically distinguish beyond age and sociodemographic characteristics, but transitions in obesity may also vary over time with cohort (Ogden, Carroll, & Flegal, 2003). For example, individuals exposed to lifestyles of the 1940s and 1950s while growing up may have different lifestyle risks than individuals born in the 1970s and 1980s. Using NHANES data and validated with data from the Nurses' Health Study and Health Professionals Follow-up Study, Wang and colleagues (Wang et al., 2007) forecast the BMI distribution in the US population based on race-, gender-, and birth cohort-specific secular trends. They found that BMI secular trends in the past three decades differ significantly by birth cohort, as well as gender and race. Their results signal a likely expansion in the population with obesity-related comorbidities as the baby boom generation approaches retirement age. However, while the more sophisticated trend models that distinguish by age, period, and cohort may be accurate, they may not be stable over time, if the underlying pattern of diet and physical activity or environmental influences on those activities shifts over time in a way not related to past trends.

While smoking is relatively simple in that it is based on a single risk factor that is relatively well identified, other risks factors are not so easily identified. For example, the consumption of alcoholic beverages is considered relatively safe if consumed in moderation and not preceding engagement in other risky behaviors (e.g., driving or jet skiing) and may even improve health. Similarly, obesity depends on the consumption of food, a necessary function for good health, as well as exercise.

Predicting the change of body weight that results from a specific change in diet or physical activity generally involves a "below the skin" mathematical model of energy balance. These models aim to link changes in diet or physical activity to changes in body weight, either at the individual or at the population level. The simplest model proposed by Hill and colleagues (Hill, Wyatt, Reed, & Peters, 2003) provides a measure of the average daily energy imbalance underlying body weight

change by assuming that a pound of weight represented 3,500 kcal excess storage and that excess energy was converted to storage at a 50 % efficiency. Based on historical trends, they estimated that the median daily energy imbalance gap between intake and expenditure needed to explain a population weight gain of 1.8–2.0 pounds per year was about 30 kcal/d. Hall and Jordan (2008) developed a spreadsheet-based model of the change in steady state body weight that explicitly included the effect of both total energy intake along with physical activity changes and metabolic adaptation impact of total energy expenditure. Both Swinburn and colleagues (Swinburn, Sacks, & Ravussin, 2009) and Hall and Chow (Guo & Hall, 2009) concluded that the increase in US food supply over the last 30 years was more than sufficient to explain the concurrent increase in weight in the US population. Their models also show that the extent to which caloric consumption will need to be reduced will depend on the initial weight of the individual and will probably involve continual reductions in daily caloric consumption, rather than a maintaining a particular reduction. The effects of caloric intake and expenditure will also depend on age. Butte and Ellis (2003) showed also that the daily energy gap was higher for children who were rapidly gaining weight compared with adults.

Environmental-Based Behavioral Models

While much of the early prevention literature focused on education and knowledge, recent literature considers the role of the environment in shaping risky behaviors. Economic factors, industry advertising, and other factors have been considered.

While smoking initiation largely takes place before age 24, cessation takes place over the life cycle of the individual and may be affected by environmental factors, such as industry advertising and general attitudes toward health. While diet and physical activity are the underlying behaviors that explain weight changes, the underlying environment can be expected to change these behaviors in similar or different ways. Understanding how the environment affects behaviors may be necessary to explain trends and changes in trends for obesity rates. A separate set of behavioral models that examine psychosocial or economic behaviors provides insight into the environmental factors that encourage obesity and, thus, set the stage for better understanding the role of policies (discussed in the next section).

Network models (Bahr, Browning, Wyatt, & Hill, 2009; Christakis & Fowler, 2007; Cohen-Cole & Fletcher, 2008; Fowler & Christakis, 2008; Halliday & Kwak, 2009; Trogden & Nonnemkaer, 2008) have been used to suggest that smoking and obesity depend on an individual's associations at home, at work, and in the geographic neighborhood. Applying a generalized linear statistical model to the Framingham longitudinal data, Christakis and Fowler (2007) shows the spread of obesity (>30 BMI) over time through social ties. Rather than estimate a statistical model, Bahr and colleagues (Bahr et al., 2009) explicitly modeled social networks using a majority rule based on the interaction with neighbors and dependent on social volatility. For a wide variety of conditions and networks, individuals with

similar BMIs were found to cluster into groups, and, if left unchecked, social forces drive these groups toward increasing obesity. They also show that many traditional weight management interventions fail because they target overweight and obese individuals without considering their surrounding cluster and wider social network. Edwards and Clarke (2009) also consider neighborhood effects but focus on a plethora of health, environment, and socioeconomic variables to understand risk factors. Their results show that social capital and poverty are strongly associated with childhood obesity.

Using an agent-based model, Hammond and Epstein (2007) examined individual behavior in the context of the physiology of dieting (and exercise) and socially influenced weight changes. They focused on the role of social influence and norms, the interaction of physiological realities with decision-making about eating, the role of exercise, and the impact of media and public health messages. They show how the core equations governing the physiology of weight change can generate many of the known facts about diet and weight gain, including the difficulty of maintaining a diet over a long period, high rates of recidivism after dieting, and substantial individual heterogeneity in the success of different types of diets.

Burke and Heiland (2007) develop a macro simulation model that examines food and nonfood consumption, focusing on the role of falling food prices in the context of a model involving endogenous body weight norms and an explicit description of human metabolism (with heterogeneity). The model accurately predicts increases in average weight and obesity rates, especially for the relative growth in upper quantile weights. A differential response to price declines across the weight distribution implies that food price affects weight more for heavier individuals. Along similar lines, Schroeter and colleagues (Schroeter, Lusk, & Tyner, 2008) developed an economic macro model to identify conditions under which price and income changes are most likely to affect weight. They found that, although raising the price of high-calorie food will decrease consumption, the resulting impact on body weight is unclear. For example, a tax on food eaten away from home could lead to an increase in body weight due to substitution of other high-caloric foods. The study shows how SMs may be used to detect unintended consequences and the importance of the scope of foods covered by a policy aimed at diet.

Models that integrate social, physiological, and economic aspects can provide deeper explanations of the observed dynamics of risk factors and suggest policies tailored to specific communities. In particular, SMs may also elucidate the implications of the interplay between genes and environment for risk factors and health policy. Research already shows that genetic traits can help set bounds on the effects that changes in the environment may have (Bouchard, 2008).

Policy Models

A substantial body of research has examined the effect of public policies on smoking attitudes and behaviors (Hopkins et al., 2001; Levy, Gitchell, &

Chaloupka, 2004; U.S. DHHS, 2000), probably more than alcohol, obesity, or illicit drug use. However, while we have good evidence on the effects of some policies (e.g., taxes and worksite restrictions), our knowledge of other policies (e.g., youth access policies, media policies, advertising bans) is relatively weak (Hopkins et al., 2001; Levy et al., 2004). Even for the better-studied policies, our knowledge is often limited in terms of how the effect of that policy depends on:

- The initial level (e.g., the amount of the tax before any policy change)
- The extent of change (e.g., the change in the tax rate or per capita media expenditures)
- The manner in which a policy is implemented (e.g., whether a clean air law is publicized and enforced)
- The period of time that it has been in effect (e.g., the effect of a policy that predominantly affects the youth will generally have increasing effects on adult prevalence with time)
- Other policies in place or currently being implemented (e.g., the effect of a public media campaign may be greater if there is already a ban on industry advertising)
- The sociodemographic or smoking characteristics of the population (e.g., a policy may have less effect on heavy than on light smokers)

Understanding these nuances is important for creating more robust tobacco control policies.

Except for a small number of empirical studies that simultaneously consider the effect of two tobacco control policies (Chaloupka & Wechsler, 1997; Farrelly, Pechacek, Thomas, & Nelson, 2003; Hu, Sung, & Keeler, 1995), most consider the effect of only one policy, thus making it difficult to learn how multiple policies interact with one another. Except for tax policy, little is known about the reactions of different sociodemographic groups (Levy et al., 2004). The other issues listed above are also not considered in most studies. Information is often not available to estimate complex statistical models; data on outcomes, policies, or other influencing factors may be poorly measured or unavailable over sufficiently long periods. In addition, the variables of interest may not exhibit sufficient variability to accurately estimate complex relationships.

In the *SimSmoke* model (Levy et al., 2010a), separate modules are developed for different types of tobacco control policies, including price interventions (taxes), clean air laws (workplace, restaurant, and other public places), mass media/comprehensive campaigns, advertising restrictions (bans/warning labels), cessation treatment policies (encouraging interventions by healthcare providers and providing financial coverage for treatments), and policies that limit youth access to cigarettes (enforcement of underage laws and vending machine and self-service bans). The policy parameters in the model used to generate the predicted effects are based on reviews of the literature (Levy et al., 2000a, 2000b; Levy & Friend, 2001, 2002a, 2002b; Levy, Friend, Holder, & Carmona, 2001) and "reasonable estimates" developed with the advice of an expert panel. In evaluating studies aimed at obesity, similar problems arise. Veerman and colleagues (Veerman, Van Beeck, Barendregt,

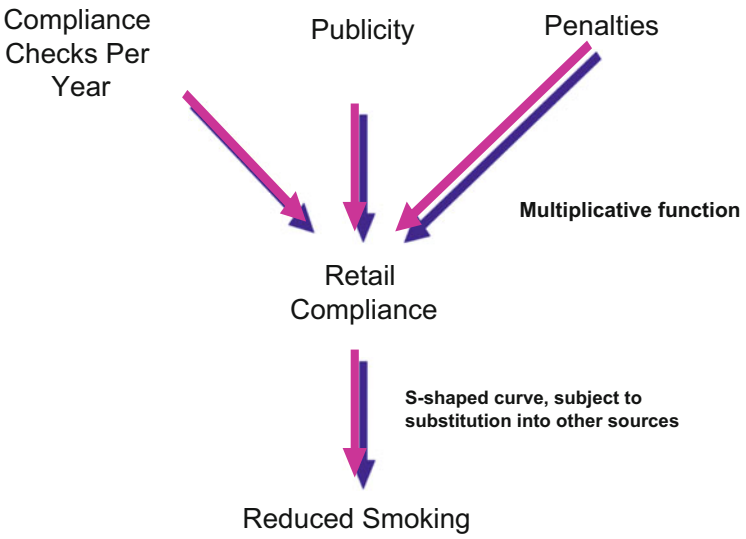

Fig. 18.4 Policy components affecting retail compliance

& Mackenbach, 2009) constructed an SM to estimate the potential effects of reducing the exposure of 6- to 12-year-old children to TV food advertising on obesity. Parameter estimates were derived from the empirical literature that relates advertising to consumption levels and consumption to body mass and from a separate Delphi study of experts' estimates of the effect of advertising on consumption.

In the *SimSmoke* model, the structure depends on theoretical constructs from the social sciences as well as empirical studies. For example, the cessation treatment module allows for substitution among different treatments and diminishing effectiveness at the population level, and the mass media module incorporates the classic S-shaped dissemination curve in which per capita expenditures must be high enough for messages to reach smokers a sufficient number of times and then shows diminishing returns. In a model of youth access policies, as illustrated in Fig. 18.4, Levy and Friend (2000) show that to effectively restrict the sales of cigarettes to minors, the law must be enforced (through compliance checks and sufficient penalty) and publicized at sufficient levels to close off most sales. Like media campaigns, expenditures on these efforts can be expected to be subject to increasing returns up to some point (to take advantage of threshold effects) and then diminishing returns. Even when a youth access policy is effectively enforced, youth may substitute cigarettes from non-purchased sources, such as older peers, parents, or theft. Other policies may be required to shut off these sources. For example, with higher cigarette taxes, older peers and parents may be less likely to smoke and less likely to provide cigarettes at the higher prices. Similarly, a policy limiting the availability of SSBs in school vending machines and cafeterias might lead to

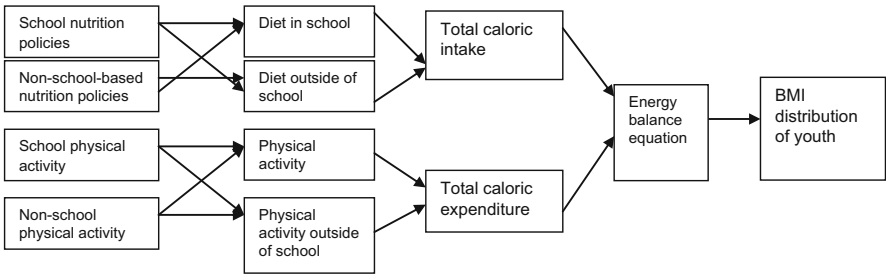

Fig. 18.5 Effect of youth-oriented policies on overweight and obesity

substitution of SSBs brought from home or more SSB consumption outside school (but see Briefel, Wilson, & Gleason, 2009; Schwartz, Novak, & Fiore, 2009).

Figure 18.5 shows the different youth-oriented obesity-oriented policies and how the effect of one policy might affect other policies. If an effective education program is implemented in school along with a policy that reduces access to SSBs, students might be less likely to consume SSBs outside school. Similarly, a nonschool policy such as restrictions on advertising SSBs or higher prices on SSBs might discourage consumption in schools of SSBs brought from home. In that case, the effects of the policies can be reinforcing. On the other hand, policies directed at consumption of the same food groups at a particular location might be less effective when combined than the sum of effects of the policies implemented alone, since they have the same purpose. Similar analysis can be applied to policies promoting physical activity (or discouraging sedentary behavior), and any interactive effects of these policies with those directed at nutrition can be considered (not shown).

Simulation models are well suited to considering the interactive effects of multiple youth-oriented policies. Similar analysis might be applied to adults in terms of considering policies in the workplace and their interactions with other nonwork-related policies. To develop a coherent and effective strategy, it will be important to consider the interactive effects with other policies in place as well as their direct effect. The *SimSmoke* model (Levy et al., 2010a) assumes that the percentage reductions from the different policies are multiplicatively applied, but also allows for some specific synergies between policies. Media campaigns are more potent when other policies are simultaneously implemented, due to reinforcing messages from their publicity. In addition, as other policies reduce the number of older smokers, youth have fewer alternative sources for acquiring cigarettes when retail sources become less accessible.

Dynamic models can also be used to show how the effects of a policy vary with the amount of time in effect. Based on evidence from prior studies, most of the policies have a large immediate effect through prevalence rates apportioned over the first 3 years that the policy is in effect. Policies continue to affect cessation and initiation rates as long as the policy is sustained. For example, a policy that provides coverage of smoking cessation treatments might be expected to have its greatest

impact in the first year as those most likely to quit and most likely to succeed take advantage of the new policy (Levy & Friend, 2002a, 2002b). Similarly, individuals most amenable to improved diet or exercise may be most likely to react in the first year that a policy is implemented and adapt over time by reverting to a poorer diet. On the other hand, the effects may build over time as the effects disseminate through social networks (Levy & Friend, 2002a, 2002b). If a policy is reversed, the effects may not be reversed (e.g., an education policy).

Dynamic models can also be useful in understanding the role of feedback loops. As individuals lose weight in response to a policy, social norms may change. The change in social norms may feed back to greater political pressure to implement new policies, thereby enhancing the effect of past policies. On the other hand, the effects of a policy could diminish over time if, for example, food manufacturers adapt their marketing practices to maintain sales of low-nutrition energy-dense foods.

One reason for implementing different policies is that they affect different demographic or smoking (e.g., heavy vs. light smokers) groups. For example, taxes have greater effects on the young and on certain minorities. Policies, such as worksite laws, have greater effects on men and those of working age due to greater labor participation and thereby have a more immediate impact on deaths attributable to smoking. Higher tobacco taxes have been found to have more prominent effects on those at younger ages (who are less addicted) and those in families with lower income (who can less afford the higher prices) (Ahmad, 2005; Farrelly & Bray, 1998; Levy & Friend, 2002a; Levy et al., 2000a). In obesity policy evaluations, studies have found that the youth at risk for overweight and from low socioeconomic status families are particularly sensitive to fruit and vegetable prices (Sturm & Datar, 2005, 2008) and that high BMI youth (Auld & Powell, 2009; Powell, Auld, Chaloupka, O'Malley, & Johnston, 2007) are more sensitive to food prices. Thus, while those at lower income may be hardest hit financially by increased prices, they might be most affected by price to improve their diet. Knowledge of the impact of the policies on particular demographic subgroups will be helpful in targeting policies to most effectively reduce future obesity prevalence.

The variation in effect of policies on those at different ages has important dynamic implications. A policy that affects youth translates into significant reductions in adult smoking rates only in later years as youth age (Levy et al., 2000a, 2000b). Similarly, obesity policies directed at youth only affect a small portion of the overall population and thus will have a limited impact on overall smoking prevalence in the near term (Levy et al., 2000b). Policies might help those at higher BMI lose weight or prevent those at risk of overweight or who have lost weight from gaining or regaining weight. While individuals can become overweight or obese at any age, policies that keep an individual from ever becoming obese may be most effective in terms of preventing relapse to regain weight that is lost. Dynamic models are useful in gauging these effects.

Ultimately, models will be needed to consider how the effect of each policy depends on the manner in which it is implemented, how that policy depends on the

other past and present policies, how the effects vary by sociodemographic group, and how the effects vary over time. Nevertheless, even when feedback loops are incorporated into the models, assumptions generally need to be made about the stability of the system. As seen in the tobacco literature (Levy et al., 2006), individual empirical studies are likely to be limited in their potential to distinguish these effects. Simulation models can combine the information from different studies to provide insights into developing a more comprehensive and effective strategy.

Conclusions

Because of the complexity of prevention problems, it is likely that different models will focus on different policies and a limited number of links from policy implementation to behaviors to risk factors to disease outcomes. Nevertheless, the different models can inform each other. The development of simulation models is itself a dynamic (iterative) process. Simulation models can provide guidance on the set of hypotheses meriting further empirical study and even point to the types of data that need to be collected. As better information becomes available, simulation models can be adapted and improved. Thereby, simulation models serve as a summary of our knowledge and a structure for improving our knowledge of prevention problems and potential solutions.

Simulation models of prevention issues, particularly policies directed at prevention, are at a nascent stage of development. Most of the models focus on one or two links in the process from changes in public policy to the health implications of obesity and often use different modeling approaches (e.g., static vs. dynamic, micro vs. macro). While the models are distinct and suited to different issues, they may complement each other in shedding light on the nature of the problem and potential solutions. As models are developed for different countries with different menus of policies and as these models are validated, the effects of policies in various combinations can be more fully understood. Nevertheless, no one model can address all questions well. Models are best designed with a specific set of questions in mind. While it is tempting to try and make models comprehensive, the model can become so unwieldy in all its complexity that the results are no longer transparent and validation becomes nearly impossible.

For the insights of the different models to be most useful, it will be important for models to be transparent. Although there are divergent views on what constitutes the optimal level of transparency (Eddy, 2006; Fendrick, 2006), it often refers to a clear statement of model structure (assumptions, equations, and algorithms), the data used to calibrate or estimate model parameters, goodness of fit to calibration data, and validation results. In particular, validation methodology is not yet as well defined as in classical statistical analyses and will need to be made explicit. Technical appendices, including supplemental material published online, can be

critical. For example, the Cancer Intervention and Surveillance Modeling Network provides online model profiles (http://cisnet.cancer.gov) with these elements.

Comparative modeling involves comparing the results of different models used to model the same problem. This allows for a more systematic examination and understanding of the specific ways that model structure, assumptions, and parameters influence model outcomes. Examples of successful comparative modeling in other health areas include the seven CISNET models for breast cancer (Berry et al., 2005; Cronin, Feuer, & Clarke, 2006) used to estimate the combined effects of screening and treatment and the Mt. Hood Challenge comparing diabetes models (Group TMHM, 2007). When the results from models differ, sensitivity analysis can help consider how their results depend on underlying assumptions and parameters. For obesity, the Robert Wood Johnson Foundation (RWJF) and the Office of Behavioral and Social Sciences Research (OBSSR) at the National Institutes of Health jointly sponsored the Collaborative Obesity Modeling Network (COMNET; http://obesitymodeling.net/), which provided support for modeling teams in the United States, the United Kingdom, Canada, and Australia to meet and compare their models. In 2009, the Comparative Modeling Network for Obesity Policy (COMPMod) was formed; it consists of six modeling teams using different modeling modalities to address a common group of policy questions. COMPMod is funded by the Eunice Kennedy Shriver National Institute on Child Health and Human Development, OBSSR, and RWJF and is part of the National Collaborative on Childhood Obesity Research (http://www.nccor.org).

As has been stated, obesity is complex, not just complicated (Finegood, Merth, & Rutter, 2010). While perhaps not as complicated, smoking, alcohol, and illicit drug policies require a broad understanding of complex phenomenon. Many factors contribute to the problem, and they relate to each other in nonlinear fashions, are subject to time delays, and change over time. Solutions will be needed that are appropriate for complex problems. Information on solutions is limited to date and resides in many different sources. Simulation models can be used to combine multiple sources of information to elucidate and test potential solutions. Thereby, more effective, comprehensive strategies may be undertaken to address complex problems.

References

Ackerman, E., Longini, I. J., Seaholms, S., & Hedin, A. (1990). Simulation of mechanisms of viral interference in influenza. *International Journal of Epidemiology, 19*(2), 444–454.

Ahmad, S. (2005). Increasing excise taxes on cigarettes in California: A dynamic simulation of health and economic impacts. *Preventive Medicine, 41*(1), 276–283.

Auld, M., & Powell, L. (2009). Economics of food energy density and adolescent body weight. *Economica, 76*(304), 719–740.

Bahr, D. B., Browning, R. C., Wyatt, H., & Hill, J. (2009). Exploiting social networks to mitigate the obesity epidemic. *Obesity (Silver Spring), 17*(4), 723–728.

Berry, D. A., Cronin, K. A., Plevritis, S., Fryback, D., Clarke, L., & Zelen, M. (2005). Effect of screening and adjuvant therapy on mortality from breast cancer. *New England Journal of Medicine, 353*(17), 1784–1792.

Beydoun, M. A., & Wang, Y. (2009). Gender-ethnic disparity in BMI and waist circumference distribution shifts in US adults. *Obesity (Silver Spring), 17*(1), 169–176.

Bibbins-Domingo, K., Coxson, P., Pletcher, M., Lightwood, J., & Goldman, L. (2007). Adolescent overweight and future adult coronary heart disease. *New England Journal of Medicine, 357* (23), 2371–2379.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceeding of the National Academy of Sciences, 99*(Suppl 3), 7280–7287.

Bouchard, C. (2008). Gene-environment interactions in the etiology of obesity: Defining the fundamentals. *Obesity (Silver Spring), 16*(Suppl 3), S5–S10.

Briefel, R. R., Wilson, A., & Gleason, P. (2009). Consumption of low-nutrient, energy-dense foods and beverages at school, home, and other locations among school lunch participants and nonparticipants. *Journal of the American Dietetic Association, 109*(2 Suppl), S79–S90.

Burke, M., & Heiland, F. (2007). Social dynamics of obesity. *Economic Inquiry, 45*(3), 571–591.

Butte, N. F., & Ellis, K. J. (2003). Comment on "Obesity and the environment: Where do we go from here?". *Science, 301*(5633), 598. author reply 598.

Cassels, S., Clark, S., & Morris, M. (2008). Mathematical models for HIV transmission dynamics: Tools for social and behavioral science research. *Journal of Acquired Immune Deficiency Syndromes, 47*(Suppl 1), S34–S39.

Chaloupka, F. J., & Wechsler, H. (1997). Price, tobacco control policies and smoking among young adults. *Journal of Health Economics, 16*(3), 359–373.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine, 357*(4), 370–379.

Cohen-Cole, E., & Fletcher, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic. *Journal of Health Economics, 27*(5), 1382–1387.

Cox, D., & Miller, H. (1965). *The theory of stochastic processes*. London: Chapman and Hall Ltd.

Cronin, K. A., Feuer, E. J., & Clarke, L. (2006). Impact of adjuvant therapy and mammography on U.S. mortality from 1975 to 2000: Comparison of mortality results from the CISNET breast cancer base case analysis. *Journal of the National Cancer Institute Monographs, 36*, 112–121.

Eddy, D. (2006). Accuracy versus transparency in harmacoeconomic modelling: Finding the right balance. *Pharmacoeconomics, 24*(9), 837–844.

Edwards, K. L., & Clarke, G. P. (2009). The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. *Social Science and Medicine, 69*(7), 1127–1134.

Farrelly, M., & Bray, J. (1998). Responses to increases in cigarette prices by race/ethnicity, income, and age groups–United States, 1976–1993. *Morbidity and Mortality Weekly Report, 47*(29), 605–609.

Farrelly, M. C., Pechacek, T. F., Thomas, K., & Nelson, D. (2003). The impact of tobacco control program expenditures on aggregate cigarette sales: 1981–2000. *Journal of Health Economics, 22*(5), 843–859.

Fendrick, A. (2006). The future of health economic modeling: Have we gone too far or not far enough? *Value in Health, 9*(3), 179–180.

Finegood, D. T., Merth, T. D., & Rutter, H. (2010). Implications of the foresight obesity system map for solutions to childhood obesity. *Obesity (Silver Spring), 18*(Suppl 1), S13–S16.

Finkelstein, E. A., Fiebelkorn, I. C., & Guijing, W. (2003). National medical spending attributable to overweight and obesity: How much, and who's paying? *Health Affairs (Millwood),* Suppl Web Exclusives: W3-219-226.

Finkelstein, E. A., Ruhm, C. J., & Kosa, K. (2005). Economic causes and consequences of obesity. *Annual Review of Public Health, 26*, 239–257.

Finkelstein, E. A., Trogdon, J. G., Brown, D., Allaire, B., Dellea, P., & Kamal-Bahl, S. (2008). The lifetime medical cost burden of overweight and obesity: Implications for obesity prevention. *Obesity (Silver Spring), 16*(8), 1843–1848.

Finkelstein, E. A., Trogdon, J. G., Cohen, J., & Dietz, W. (2009). Annual medical spending attributable to obesity: Payer- and service-specific estimates. *Health Affairs (Millwood), 28*(5), w822–w831.

Flanagan, W. M., Le Petit, C., Berthelot, J., White, K., Coombs, B., & Jones-McLean, E. (2003). Potential impact of population-based colorectal cancer screening in Canada. *Chronic Diseases in Canada, 24*(4), 81–88.

Flegal, K. M. (2007). The effects of changes in smoking prevalence on obesity prevalence in the United States. *American Journal of Public Health, 97*(8), 1510–1514.

Flegal, K. M., & Graubard, B. I. (2009). Estimates of excess deaths associated with body mass index and other anthropometric variables. *American Journal of Clinical Nutrition, 89*(4), 1213–1219.

Flegal, K. M., Graubard, B. I., Williamson, D., & Gail, M. (2007a). Cause-specific excess deaths associated with underweight, overweight, and obesity. *JAMA, 298*(17), 2028–2037.

Flegal, K. M., Graubard, B. I., Williamson, D., & Gail, M. (2007b). Impact of smoking and preexisting illness on estimates of the fractions of deaths associated with underweight, overweight, and obesity in the US population. *American Journal of Epidemiology, 166*(8), 975–982.

Flegal, K. M., Williamson, D. F., Pamuk, E., & Rosenberg, H. (2004). Estimating deaths attributable to obesity in the United States. *American Journal of Public Health, 94*(9), 1486–1489.

Fontaine, K. R., Redden, D. T., Wang, C., Westfall, A., & Allison, D. (2003). Years of life lost due to obesity. *JAMA, 289*(2), 187–193.

Fowler, J. H., & Christakis, N. A. (2008). Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais. *Journal of Health Economics, 27*(5), 1400–1405.

Goldman, D. P., Zheng, Y., Girosi, F., Michaud, P., Olshansky, S., Cutler, D., et al. (2009). The benefits of risk factor prevention in Americans aged 51 years and older. *American Journal of Public Health, 99*(11), 2096–2101.

Group TMHM. (2007). Computer modeling of diabetes and its complications: A report on the Fourth Mount Hood Challenge Meeting. *Diabetes Care, 30*(6), 1638–1646.

Guo, J., & Hall, K. D. (2009). Estimating the continuous-time dynamics of energy and fat metabolism in mice. *PLoS Computational Biology, 5*(9), e1000511.

Hall, K. D., & Jordan, P. N. (2008). Modeling weight-loss maintenance to help prevent body weight regain. *American Journal of Clinical Nutrition, 88*(6), 1495–1503.

Halliday, J., & Kwak, M. (2009). Weight gain in adolescent peers. *Economics and Human Biology, 7*(2), 181–190.

Hammond, D., & Epstein, J. (2007). *Exploring price-independent mechanisms in the obesity epidemic (paper no. 48)*. Washington, DC: The Brookings Institution.

Hill, J. O., Wyatt, H. R., Reed, G., & Peters, J. (2003). Obesity and the environment: Where do we go from here? *Science, 299*(5608), 853–855.

Homer, J. B., & Hirsch, G. B. (2006). System dynamics modeling for public health: Background and opportunities. *American Journal of Public Health, 96*(3), 452–458.

Hopkins, D. P., Briss, P. A., Ricard, C., Husten, C., Carande-Kulis, V., Fielding, J., et al. (2001). Reviews of evidence regarding interventions to reduce tobacco use and exposure to environmental tobacco smoke. *American Journal of Preventive Medicine, 20*(Suppl 2), 16–66.

Hu, T. W., Sung, H. Y., & Keeler, T. (1995). Reducing cigarette consumption in California: Tobacco taxes vs an anti-smoking media campaign. *American Journal of Public Health, 85*(9), 1218–1222.

Huang, E. S., Basu, A., O'Grady, M., & Capretta, J. (2009). Projecting the future diabetes population size and related costs for the U.S. *Diabetes Care, 32*(12), 2225–2229.

Huang, T. T., & Glass, T. A. (2008). Transforming research strategies for understanding and preventing obesity. *JAMA, 300*(15), 1811–1813.

Jia, H., & Lubetkin, E. I. (2009). The statewide burden of obesity, smoking, low income and chronic diseases in the United States. *Journal of Public Health (Oxford), 31*(4), 496–505.

Jones, A. P., Homer, J. B., Murphy, D., Essien, J., Milstein, B., & Seville, D. (2006). Understanding diabetes population dynamics through simulation modeling and experimentation. *American Journal of Public Health, 96*(3), 488–494.

Kopec, J. A., Sayre, E. C., Flanagan, W. M., Fines, P., Cibere, J., Rahman, M. M., Bansback, N. J., Anis, A. H., Jordan, J. M., Sobolev, B., Aghajanian, J., Kang, W., Greidanus, N. V., Garbuz, D. S., Hawker, G. A., Badley, E. M. (2010). Development of a population-based microsimulation model of osteoarthritis in Canada. *Osteoarthritis Cartilage*, *18*(3), 303–311. doi: 10.1016/j.joca.2009.10.010. Epub 2009 Oct 23.PMID:19879999.

Lemp, J., McWethy, L., & Kockelman, K. (2007). From aggregate methods to microsimulation: Assessing benefits 80 of microscopic activity-based models of travel demand. *Transportation Research Record, 1994*, 80–88.

Levy, D. T., Bauer, J. E., & Lee, H. (2006). Simulation modeling and tobacco control: Creating more robust public health policies. *American Journal of Public Health, 96*(3), 494–498.

Levy, D. T., Chaloupka, F., Gitchell, J., Mendez, D., & Warner, K. E. (2002). The use of simulation models for the surveillance, justification and understanding of tobacco control policies. *Health Care Management Science, 5*(2), 113–120.

Levy, D. T., Cummings, K. M., & Hyland, A. (2000a). Increasing taxes as a strategy to reduce cigarette use and deaths: Results of a simulation model. *Preventive Medicine, 31*(3), 279–286.

Levy, D. T., Cummings, K. M., & Hyland, A. (2000b). A simulation of the effects of youth initiation policies on overall cigarette use. *American Journal of Public Health, 90*(8), 1311–1314.

Levy, D. T., & Friend, K. B. (2000). A simulation model of tobacco youth access policies. *Journal of Health Politics, Policy and Law, 25*(6), 1023–1050.

Levy, D. T., & Friend, K. (2001). A computer simulation model of mass media interventions directed at tobacco use. *Preventive Medicine, 32*(3), 284–294.

Levy, D. T., & Friend, K. (2002a). Examining the effects of tobacco treatment policies on smoking rates and smoking related deaths using the SimSmoke computer simulation model. *Tobacco Control, 11*(1), 47–54.

Levy, D. T., & Friend, K. (2002b). A simulation model of policies directed at treating tobacco use and dependence. *Medical Decision Making, 22*(1), 6–17.

Levy, D. T., Friend, K., Holder, H., & Carmona, M. (2001). Effect of policies directed at youth access to smoking: Results from the SimSmoke computer simulation model. *Tobacco Control, 10*(2), 108–116.

Levy, D. T., Gitchell, J. G., & Chaloupka, F. (2004). The effects of tobacco control policies on smoking rates: A tobacco control scorecard. *Journal of Public Health Management and Practice, 10*, 338–351.

Levy, D., Mabry, P., Graham, A., Orleans, T., & Abrams, D. (2010a). Reaching Healthy People 2010 by 2013: A SimSmoke simulation. *American Journal of Preventive Medicine, 38*(38), S373–S381.

Levy, D., Mabry, P., Graham, A., Orleans, T., & Abrams, D. (2010b). Exploring scenarios to dramatically reduce smoking prevalence: A simulation model of the three-part cessation process. *American Journal of Public Health, 100*(7), 1253–1259.

Levy, D. T., Nikolayev, N., Mumford, E., & Compton, C. (2005). The Healthy People 2010 smoking prevalence and tobacco control objectives: Results from the SimSmoke Tobacco Control Policy Simulation Model. *Cancer Causes and Control, 16*(4), 359–371.

Levy, D., Tworek, C., Hahn, E., & Davis, R. (2008). The Kentucky SimSmoke Tobacco Policy Simulation Model: Reaching Healthy People 2010 goals through policy change. *Southern Medical Journal, 101*(5), 503–507.

Lightwood, J., Bibbins-Domingo, K., Coxson, P., Wang, C., Williams, L., & Goldman, L. (2009). Forecasting the future economic burden of current adolescent overweight: An estimate of the coronary heart disease policy model. *American Journal of Public Health, 99*(12), 2230–2237.

Mendez, D., & Warner, K. E. (2004). Adult cigarette smoking prevalence: Declining as expected (not as desired). *American Journal of Public Health, 94*(2), 251–252.

Mendez, D., Warner, K. E., & Courant, R. N. (1998). Has smoking cessation ceased? Expected trends in the prevalence of smoking in the United States. *American Journal of Epidemiology, 148*(3), 249–258.

Miller, T. R., Levy, D. T., Cohen, M., & Cox, K. (2006). Costs of alcohol and drug-involved crime. *Prevention Science, 7*(4), 333–342.

Milstein, B., Jones, A., Homer, J., Murphy, D., Eissen, J., & Seville, D. (2007). Charting plausible futures for diabetes prevalence in the United States: A role for system dynamics simulation modeling. *Preventing Chronic Disease, 4*(3), A52.

Ogden, C. L., Carroll, M. D., & Flegal, K. (2003). Epidemiologic trends in overweight and obesity. *Endocrinology and Metabolism Clinics of North America, 32*(4), 741–760. vii.

Orcutt, G. (1957). A new type of socio-economic system. *Review of Economics and Statistics, 80*, 1081–1100.

Orme, M. E., Hogue, S. L., Kennedy, L. M., Paine, A. C., & Godfrey, C. (2001). Development of the health and economic consequences of smoking interactive model. *Tobacco Control, 10*(1), 55–61.

Peeters, A., Barendregt, J. J., Willekens, F., Mackenbach, J., Al Mamun, A., Bonneux, L., et al. (2003). Obesity in adulthood and its consequences for life expectancy: A life-table analysis. *Annals of Internal Medicine, 138*(1), 24–32.

Popkin, B. M., Kim, S., Rusev, E., & Zizza, A. (2006). Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases. *Obesity Reviews, 7*(3), 271–293.

Powell, L., Auld, M., Chaloupka, F., O'Malley, P., & Johnston, L. (2007). Access to fast food and food prices: Relationship with fruit and vegetable consumption and overweight among adolescents. *Advances in Health Economics and Health Services Research, 17*, 23–48.

Roberts, E. B., Homer, J., Kasabian, A., & Varrell, M. (1982). A systems view of the smoking problem: Perspective and limitations of the role of science in decision-making. *International Journal of Bio-Medical Computing, 13*(1), 69–86.

Rodgers, A., Ezzati, M., Vander Hoorn, S., Lopez, A., Lin, R., Murray, C., et al. (2004). Distribution of major health risks: Findings from the Global Burden of Disease study. *PLoS Med, 1* (1), e27.

Schroeter, C., Lusk, J., & Tyner, W. (2008). Determining the impact of food price and income changes on body weight. *Journal of Health Economics, 27*, 45–68.

Schwartz, M. B., Novak, S. A., & Fiore, S. (2009). The impact of removing snacks of low nutritional value from middle schools. *Health Education and Behavior, 36*(6), 999–1011.

Sonnenberg, F. A., & Beck, J. R. (1993). Markov models in medical decision making: A practical guide. *Medical Decision Making, 13*(4), 322–338.

Sterman, J. D. (2006). Learning from evidence in a complex world. *American Journal of Public Health, 96*(3), 505–514.

Stewart, S. T., Cutler, D. M., & Rosen, A. (2009). Forecasting the effects of obesity and smoking on U.S. life expectancy. *New England Journal of Medicine, 361*(23), 2252–2260.

Sturm, R., & Datar, A. (2005). Body mass index in elementary school children, metropolitan area food prices and food outlet density. *Public Health, 119*(12), 1059–1068.

Sturm, R., & Datar, A. (2008). Food prices and weight gain during elementary school: 5-year update. *Public Health, 122*(11), 1140–1143.

Swinburn, B., Gill, T., & Kamanyika, S. (2005). Obesity prevention: A proposed framework for translating evidence into action. *Obesity Reviews, 6*(1), 23–33.

Swinburn, B., Sacks, G., & Ravussin, E. (2009). Increased food energy supply is more than sufficient to explain the US epidemic of obesity. *American Journal of Clinical Nutrition, 90* (6), 1453–1456.

Tengs, T. O., Ahmad, S., Moore, R., & Gage, E. (2004). Federal policy mandating safer cigarettes: A hypothetical simulation of the anticipated population health gains or losses. *Journal of Policy Analysis and Management, 23*(4), 857–872.

Tengs, T. O., Osgood, N. D., & Lin, T. (2001). Public health impact of changes in smoking behavior: Results from the Tobacco Policy Model. *Medical Care, 39*(10), 1131–1141.

Trogden, J., & Nonnemkaer, J. (2008). Peer effects in adolescent overweight. *Journal of Health Economics, 27*(5), 1378–1387.

Trogdon, J. G., Finkelstein, E. A., Hylands, T., Dellea, P., & Kamal-Bahl, S. (2008). Indirect costs of obesity: A review of the current literature. *Obesity Reviews, 9*(5), 489–500.

U.S. DHHS. (2000). *Reducing tobacco use: A report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

van Baal, P. H., Polder, J. J., de Wit, G., Hoogenveen, R., Feenstra, T., Boshuizen, H., et al. (2008). Lifetime medical costs of obesity: Prevention no cure for increasing health expenditure. *PLoS Med, 5*(2), e29.

Van Meijgaard, J., Fielding, J. E., & Kominski, G. (2009). Assessing and forecasting population health: Integrating knowledge and beliefs in a comprehensive framework. *Public Health Reports, 124*(6), 778–789.

Veerman, J. L., Van Beeck, E. F., Barendregt, J., & Mackenbach, J. (2009). By how much would limiting TV food advertising reduce childhood obesity? *European Journal of Public Health, 19*(4), 365–369.

Wang, Y., & Beydoun, M. A. (2007). The obesity epidemic in the United States–gender, age, socioeconomic, racial/ethnic, and geographic characteristics: A systematic review and meta-regression analysis. *Epidemiologic Reviews, 29*, 6–28.

Wang, Y. C., Colditz, G. A., & Kuntz, K. (2007). Forecasting the obesity epidemic in the aging U.S. population. *Obesity (Silver Spring), 15*(11), 2855–2865.

Wolfson, M. C. (1994). POHEM—A framework for understanding and modelling the health of human populations. *World Health Statistics Quarterly, 47*(3–4), 157–176.

Chapter 19
Meta-analysis in Prevention Science

Sandra Jo Wilson and Emily E. Tanner-Smith

Introduction

This chapter describes the fundamental elements of meta-analysis, with particular emphasis on its relevance to prevention science. The goal is to provide readers with a basic understanding of what a meta-analysis is, how to identify meta-analysis topics appropriate to prevention science, how to interpret results from meta-analysis, and how to identify some of the potential biases in meta-analysis; in short, our goal is to create intelligent consumers of meta-analysis. Armed with knowledge about some of the common ways that meta-analytic techniques can be used in prevention science research, we encourage readers interested in conducting a meta-analysis to seek more comprehensive resources on the statistical methods unique to this form of research (e.g., Borenstein, Hedges, Higgins, & Rothstein, 2009; Cooper, Hedges, & Valentine, 2009; Lipsey & Wilson, 2001).

What Is Meta-analysis?

The term "meta-analysis," coined by Glass (1976), encompasses a range of techniques for quantitatively collecting data from a preselected set of primary research studies and applying specialized statistical analyses that synthesize findings across studies. In contrast to primary studies that may use individuals, families, classrooms, or schools as the unit of analysis, it is the primary studies themselves that are the unit of analysis in a meta-analysis. Because different primary studies often use different measures to represent the same underlying constructs, meta-analysts are tasked with standardizing findings across studies to make results comparable.

S.J. Wilson (✉)
Peabody Research Institute, Vanderbilt University, Peabody #181, 230 Appleton Place, Nashville, TN 37203, USA
e-mail: sandra.j.wilson@vanderbilt.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_19,

This is achieved by calculating *effect sizes* from the data reported in each primary study. An effect size is a standardized, quantitative index representing the magnitude and direction of a relationship. By representing the findings of each study included in a meta-analysis in the same form, the effect size permits a synthesis of those findings across studies. There are a wide variety of effect size indices that can be meta-analyzed, including those representing measures of central tendency (means, proportions), pretest–posttest contrasts, group contrasts (mean differences), and associations between two variables (correlations, odds ratios); the effect size index chosen for a given meta-analysis will depend on both the goals of the meta-analysis and the types of statistical information commonly reported in the primary research literature of interest.

Meta-analysis is closely related to, and often overlaps with, another form of research synthesis called a systematic review. A systematic review seeks to identify, collate, and systematically summarize all empirical evidence on a specific research topic, using explicit, systematic, transparent, replicable methods that are designed to minimize bias. This is in contrast to traditional narrative reviews, which often lack explicit eligibility criteria and transparency and tend to summarize the results of included studies in a subjective manner. Like systematic reviews, meta-analyses also seek to identify, collate, and summarize empirical evidence, but they use statistical methods to produce a quantitative summary of the results of a set of studies. While many systematic reviews use meta-analysis as the method of statistically synthesizing the included studies' findings, not all systematic reviews will include a meta-analysis (e.g., they may only provide a narrative summary of the studies). Meta-analyses are not always based on systematic reviews and may instead be based on unsystematic or non-exhaustive searches of the literature. Most meta-analyses, however, endeavor to be exhaustive in their search so that results may generalize to a broader population of studies.

Rationale for Meta-analysis

Several features of meta-analysis make it particularly applicable to prevention science. First, by coding detailed information about the characteristics of the included studies into a database that can be analyzed statistically, meta-analysis provides an organized, systematic, and comprehensive approach to representing the nature and findings of numerous complex and diverse individual studies (Cooper, 1984). Second, the statistical aggregation across multiple respondent samples involved in meta-analysis yields more statistically reliable estimates than individual studies and, in particular, inhibits misinterpretation of sampling error as real differences among studies (Hunter, Schmidt, & Jackson, 1982; Schmidt, 1992). However, where there are real differences in findings among studies, meta-analysis permits statistically sophisticated analysis of the source of those differences. Statistical models tailored for meta-analytic data can reveal common findings across studies despite differences in study method and procedure that may obscure the real

relationships of interest (Cooper et al., 2009; Hunter & Schmidt, 1990). Finally, by dealing with the full range of samples, variables, and relationships in a body of research, meta-analysis can potentially present a synthesis of empirical findings on a topic that has more scope, depth, and generality than any one primary study can provide. Indeed, single studies sometimes have an enormous impact on public policy, without adequate consideration given to the possibility of sampling error as well as limitations of generalizability. Meta-analysis provides a set of techniques for summarizing a body of research in ways that can provide comprehensive policy-relevant information that is more reliable and defensible than results from any primary study alone.

Of course, meta-analysis requires a body of empirical research to be on hand for synthesis, which means that meta-analysis can be applied only when there is sufficient primary research to review. It also requires that the research to be synthesized be of sufficient similarity to be compared and that the research yield quantitative results that can be represented as effect sizes. Thus, while meta-analysis can provide an additional tool in the arsenal of research methods applicable to prevention scientists, it cannot stand alone and requires primary research for its existence. We now turn to a discussion of the basic phases or steps in a meta-analysis. As with any type of research, there are potential threats to the validity of the inferences that can be drawn from meta-analysis. As we outline the steps in a meta-analysis, we will weave in a discussion of some of the more common sources of bias that represent threats to validity in meta-analysis (for more detailed discussions see Cooper, 2010; Matt & Cook, 2009). Bias refers to a systematic error in findings, in contrast to the imprecision associated with sampling error inherent in all inferential statistics. It is important to note that traditional narrative reviews are subject to many of the same biases as meta-analysis; in fact, narrative reviews are generally at greater risk for bias because of the subjective way in which studies are included and interpreted. Fortunately, meta-analysts can minimize most of the biases we discuss below through careful planning at all stages of the meta-analysis. Meta-analysts can also empirically assess the possibility of some sources of bias, something not possible with narrative reviews.

Basic Elements of a Meta-analysis

Developing a Research Question

While meta-analyses can differ widely in scope and purpose, all generally involve the same basic steps, the first of which is developing a research question. Three broad types of research questions are particularly appropriate for meta-analysis in the field of prevention science: (1) questions about the etiology or epidemiology of particular social problems; (2) questions about the efficacy or effectiveness of interventions for solving those social problems; and (3) questions about group

differences, either between naturally occurring groups (i.e., males and females) or between groups defined by researchers (e.g., between different diagnostic groups).

One of the pillars of prevention science research involves understanding the etiology and epidemiology of social problems, and one form of meta-analysis is well-suited to such questions of etiology—meta-analyses of correlational relationships between risk and protective factors and concurrent or later problems. Such meta-analyses can form the foundation for a risk-reduction/protection-enhancement approach to prevention (Battin-Pearson et al., 2000; Kraemer et al., 1997; Mrazek & Haggerty, 1994) by seeking to identify the significant predictors of the problem of interest. For example, one recent meta-analysis synthesized longitudinal correlation effect sizes from 41 studies to examine the predictive strength of risk factors for later delinquent and criminal behavior among children and adolescents, focusing on the differential strength of risk factors across developmental life stages and domains of risk (Tanner-Smith, Wilson, & Lipsey, 2011). One of the conclusions from the study was that family risk factors (e.g., harsh parenting, low family cohesion) occurring during childhood were strong risk factors for later adolescent crime and delinquency. Meta-analyses such as these are thus particularly useful for prevention scientists and can support two central elements of prevention programming: (1) identification and selection of individuals or groups of sufficiently high risk to be appropriate for services and/or (2) development of service programs to ameliorate targeted risk factors and/or enhance selected protective factors with the expectation that this will prevent, or at least mitigate, problematic outcomes or improve positive ones (Catalano, Hawkins, Berglund, Pollard, & Arthur, 2002; Kellam, Koretz, & Mościcki, 1999).

The second category of research questions particularly relevant to prevention scientists is questions about intervention effectiveness, and meta-analyses focusing on intervention effectiveness are quite common in the field. The Campbell Collaboration (http://www.campbellcollaboration.org/) publishes systematic reviews and meta-analyses of intervention effects, many of which focus on prevention and intervention programs. For example, one such meta-analysis synthesized results from 73 studies to examine the effects of universal school-based social information-processing interventions on school-age children's aggressive and disruptive behavior; the authors found an overall beneficial effect of the interventions, although the magnitude of effect was somewhat small and varied across different levels of treatment dosage (Wilson & Lipsey, 2006). Another Campbell Collaboration review synthesized results from 23 studies to examine the effects of exercise on children and youth's self-esteem, with results indicating overall beneficial effects on self-esteem outcomes (Ekeland, Heian, Hagen, Abbott, & Nordheim, 2005). Intervention meta-analyses can also be important for identifying harmful intervention approaches, as in the case of a meta-analysis of nine studies that examined whether Scared Straight juvenile awareness programs were effective for preventing juvenile delinquency; in that meta-analysis, the results indicated that participation in juvenile awareness programs was associated with higher levels of juvenile delinquency (Petrosino, Petrosino, & Buehler, 2004). Meta-analyses focusing on questions of intervention effectiveness are clearly applicable to prevention science

because they can identify effective program strategies, reveal ineffective or harmful strategies, and, as we will describe in more detail below, examine the conditions under which programs may be more or less effective.

The third major class of research questions appropriate for meta-analysis and of interest to prevention scientists has to do with group differences. Understanding the nature of differences between males and females on math achievement or studying the achievement gap for minority (or low socioeconomic status) students versus majority students is critical for understanding the nature of social problems and can inform the design of interventions that might be appropriate for different subgroups of individuals. For instance, one recent meta-analysis (Lindberg, Hyde, Linn, & Petersen, 2010) synthesized results from 242 studies that examined gender differences in mathematics achievement. The authors concluded that there was not a significant gender gap in mathematics achievement among school-age youth, although there was some evidence of a gender gap in math achievement among high school and college samples (vs. elementary and middle schools). Meta-analyses addressing research questions of group differences are particularly relevant for understanding whether certain groups of individuals may or may not need targeted prevention programs. Another application of group differences meta-analysis involves comparing groups of individuals created by the researchers, such as studies comparing outcomes or symptomatology for attention deficit hyperactive disorder patients to normal individuals (e.g., Bálint et al., 2009).

Potential Bias at the Research Question Stage

When developing the research questions of interest for a meta-analysis, an important issue to consider is whether the types of primary studies to be included are similar enough to be synthesized in a single meta-analysis. Meta-analyses that include diverse types of primary studies in a single analysis have been criticized for mixing "apples and oranges," which critics surmise to yield meaningless results. Robert Rosenthal once responded to this critique by stating that in some cases, a meta-analyst may be interested in neither apples nor oranges, but fruit salad (as cited in Borenstein et al., 2009). In the end, decisions about the breadth of studies to include in a meta-analysis depend on (and will vary according to) the stated goals of the project.

For instance, a researcher may be interested in evaluating the effectiveness of the Olweus Bullying Prevention Program and, therefore, decides to synthesize results from all randomized controlled studies that compared intervention programs using the Olweus Bullying Prevention Program manual with some control condition. To draw conclusions about this specific program's overall effectiveness, it would not make sense for this meta-analysis to also include studies of other bullying prevention programs that did not follow the Olweus manual. In contrast, if a researcher was interested in the comparative effectiveness of different types of bullying prevention programs (i.e., interested in fruit salad), they could justifiably include studies using a wide variety of bullying prevention programs (e.g., Lions Quest,

Olweus, Ripple Effects, non-manual-based bullying programs). The authors of such a meta-analysis might then empirically examine the comparative effectiveness of those programs through moderator analysis, as discussed later in this chapter. As such, it is always important to consider the breadth of studies included in a meta-analysis and whether the study results can be synthesized in a way that allows inference to a meaningful population given the research question(s) of interest.

Defining Eligibility Criteria

The next stage of a meta-analysis is to define eligibility criteria, which follow directly from the research questions. Eligibility criteria should explicitly define the types of studies that are eligible for inclusion and give the reader a clear idea of the nature of the literature being reviewed. The specifics of the eligibility criteria will vary depending on the research question, but they will generally include four primary components and several secondary components.

The Topic

Most importantly, the eligibility criteria must identify the distinguishing features of the research to be included. If the research question is one of etiology and, for example, involves studying the predictors of later antisocial behavior, the criteria should define the predictors and outcomes of interest, how they are put into operation, and the types of relationships between those variables that are eligible. If the research question focuses on the effects of an intervention, the eligibility criteria must define the critical features of that intervention so that the intervention of interest can be distinguished from other types of interventions (which may or may not be similar). Research questions on group differences necessitate eligibility criteria that demarcate the boundaries for the groups of interest and the comparisons between those groups that are relevant.

The Population

The types of research participants that are relevant must also be specified in the eligibility criteria. Most commonly, criteria with regard to the population of interest specify the gender, ethnic, socioeconomic, and age groups that are included (or excluded). When age or socioeconomic criteria are specified, the boundaries must be specific and clearly defined. For example, if middle school children are the pertinent population, the eligibility criteria should specify what age and/or grade levels constitute middle school students and how cases where middle school students are mixed in with elementary or high school students are to be handled. In addition, whether specialized populations are eligible or not should be clearly

specified. For prevention science researchers focused on educational research, for example, the eligibility criteria should clearly specify whether special education, learning disabled, or other special populations are considered eligible for the review.

Pertinent Variables

The variables of interest to the researcher must also be clearly identified. For example, if the meta-analysis is focused on intervention effects, the specific outcomes of the intervention that are eligible should be specified. Criteria here would also include limitations on the timing of measurement of the outcomes, as well as any restrictions with regard to the source and nature of the measurement. For questions about etiology, specifying the eligible pertinent variables can overlap with eligibility criteria on the topic of the meta-analysis, but further clarifying information about the measurement and timing characteristics can be provided when specifying the pertinent variables. In addition, criteria with regard to the statistical findings of the primary studies would be relevant here. Primary studies need to provide sufficient quantitative information to compute effect sizes. In addition, some meta-analysts may be interested only in certain forms of data, such as continuous or binary outcomes.

Research Methods

Finally, the eligible research methods must be specified. When studying questions about intervention effects, the types of experimental or quasi-experimental designs that are relevant should be clearly specified and excluded designs enumerated. For some types of interventions, pretest–posttest designs, time series, or single-subject designs might also be relevant for any given analysis. The research methodologies corresponding to a focus on etiology may use longitudinal or cross-sectional designs or both and report on the statistical associations among variables. Research methodologies for questions of group differences would include details about how contrived groups are constituted or the types of comparisons made between groups that are considered pertinent.

Secondary Criteria

Other secondary eligibility criteria may involve the cultural and linguistic range of the studies included, the timeframe of the literature, and the publication types that are to be included or excluded.

The PICOS Framework

We have described eligibility criteria for meta-analysis somewhat generically thus far, in a way that is applicable to all forms of meta-analysis. There is, however, a useful framework for developing eligibility criteria for intervention meta-analyses that deserves mention. This framework includes five primary components and is described with the acronym PICOS (population, intervention, comparison, outcome, study design; Higgins & Green, 2011). Under the PICOS framework, the eligibility criteria should specify the types of research participants in the primary studies of interest (population); the critical features of the intervention under study, as well as its dose, format, frequency, duration, timing, and so forth (intervention); the types of comparison conditions that are eligible, whether no treatment, treatment as usual, placebo, or some other type of intervention (comparison); the outcome constructs of interest, including the timing of measurement, operationalization, and source (outcome); and, finally, the types of study designs eligible for inclusion such as randomized, nonrandomized, pretest–posttest only, and so forth (study design).

Potential Bias at the Eligibility Criteria Stage

At the eligibility criteria stage, the meta-analyst is tasked with creating clearly defined criteria that identify the characteristics of the studies eligible for inclusion in the meta-analysis. At this stage of a meta-analysis, bias may occur due to any ambiguity in the operational definitions of the key constructs of interest specified in each of the eligibility criteria (e.g., population, intervention, outcomes, research designs). Consider a researcher proposing to conduct a meta-analysis on the effects of cognitive behavioral therapy (CBT) interventions on middle school students' tobacco use. The researcher could naively rely on study authors' descriptions of the interventions and make eligible only those studies in which programs were explicitly labeled as CBT. If the researcher is truly interested in the effectiveness of programs that use the components of CBT, reliance on a program label could result in the inclusion of studies reporting implementing "CBT" when they did not, in fact, include any elements of CBT. Moreover, such reliance on a program label could also result in the exclusion of studies that actually did use cognitive behavioral principles but were not specifically billed as CBT (like a social problem-solving intervention). In this case, the population of intervention types to which results could be generalized would be somewhat ambiguous and may not provide an adequate answer to the research question.

Another type of construct ambiguity common in intervention meta-analyses is lack of specificity about the nature of the comparison groups, especially those labeled as "practice as usual." Without a clear definition of the boundaries of "practice as usual," it is difficult to assess the type and level of services received by comparison group participants, making interpretation of the resulting effect sizes

ambiguous. Indeed, it is difficult to make statements about effective interventions when the type of service received by the comparator groups is ambiguous or undefined. Similar threats to validity can occur with all relevant study constructs (e.g., populations, pertinent variables, study designs). Regardless of the goals of meta-analysis, conceptual ambiguity at the research question and eligibility criteria stages can affect later inferences and always deserves careful consideration.

Literature Search

Using the eligibility criteria as a guide, meta-analysis next involves conducting and documenting a systematic search for all studies that meet the eligibility criteria. Ideally, the search should involve transparent, diverse, and iterative procedures to locate the population of studies relevant to the meta-analysis. Excellent resources are available on conducting literature searches, but the distinguishing feature of most searches for meta-analysis (and even more so for exhaustive systematic reviews) is that they involve multiple sources. The sources commonly include electronic citation databases that house published reports (e.g., ERIC, MEDLINE, PsycINFO), but also sources for unpublished (or "gray") literature (Rothstein & Hopewell, 2008). Internet searches, hand searches of key journals, contact with experts in the field, and reference harvesting from previous meta-analyses, systematic reviews, and narrative literature reviews are commonly used to identify gray literature that may not be indexed in standard electronic bibliographic databases.

Potential Bias at the Literature Search Stage

When conducting a meta-analysis, well-executed literature searches are those that are broad in scope, diverse in sources, and transparent in methods with the ultimate goal of minimizing the potential for publication bias. Publication bias is one of many types of reporting or dissemination biases that have the potential to influence the validity of generalizations from meta-analysis. It is due primarily to the fact that studies with large effects are more likely to be published than those with small or null effects (Rothstein, Sutton, & Borenstein, 2005; Sterne, Egger, & Moher, 2008). This has often been referred to as the "filedrawer" problem—that "journals are filled with the 5 % of studies that show Type I errors, while the file drawers back at the lab are filled with the 95 % of the studies that show non-significant (e.g., $p < 0.05$) results" (Rosenthal, 1979). There is, indeed, a large body of empirical literature documenting that primary studies with statistically significant and/or "positive" effects (i.e., those in the hypothesized direction) are more likely to be submitted for publication (Cooper, DeNeve, & Charlton, 1997; Dickersin, 1997). And in intervention meta-analyses, published studies do tend to yield larger treatment effect estimates compared with unpublished studies (Lipsey & Wilson, 1993).

When a meta-analysis includes only published research, there is the potential that overall effects may be inflated as a result.

There are a range of other reporting biases, all of which increase the likelihood of overestimating true study results. For instance, time-lag bias results from studies with large effects being published faster than those with small effects (Hopewell & Clarke, 2001). Multiple publication bias results when studies with large effects are published in multiple reports and are, therefore, more easily identified in searches (Reyes, Panza, Martin, & Bloch, 2011). Location bias occurs when large effect sizes are published in more easily accessible locations (Pittler, Abbot, Harkness, & Ernst, 2000). Citation bias is the result of studies with large effects being cited more often in other publications, making them more easily identifiable through reference harvesting (Gotzsche, 1987). Language bias is the result of studies with large effects being more likely to be published in English (Egger et al., 1997). Finally, outcome reporting bias occurs when primary study authors selectively report results for only those outcomes that show large or significant effects (Tannock, 1996).

Any type of reporting bias can distort the results of a meta-analysis. There are several instances in the field of medicine where meta-analysts would have reached different conclusions about the effectiveness of treatments for ovarian cancer, heart disease, or thyroid disease depending on whether results from unpublished studies were included in the meta-analyses (Chalmers, 2001; Rennie, 1997; Simes, 1986). To minimize the possibility of publication bias, most meta-analysts, therefore, attempt to identify this difficult-to-locate "gray literature" by using diverse search strategies other than standard electronic bibliographic databases such as PsycINFO or MEDLINE. Meta-analysts should also assess the possibility of publication bias through the use of exploratory statistical procedures (see Rothstein et al., 2005). Although no statistical tests can definitively answer "Does this meta-analysis suffer from publication bias?", it is important for meta-analysts to use the exploratory tools currently available to at least acknowledge whether findings are at risk of such bias. Furthermore, consumers of meta-analysis should pay attention to how literature search procedures and publication bias analyses are reported.

Data Extraction or Study Coding

Once the set of eligible studies is identified and obtained, a meta-analysis uses objective and systematic coding procedures to extract information from the eligible studies. Meta-analysis must attend to variation in results across studies and attempt to distinguish variation attributable to systematic differences among studies from variation attributable to sampling error and other unsystematic sources. Furthermore, it is also important to distinguish whether systematic differences across studies are due to substantive, methodological, or procedural differences in those studies. Therefore, in addition to extracting the effect sizes that index each study's findings, other types of information are also relevant. The types of information extracted from studies in a meta-analysis vary across the different types of meta-analysis,

but generally include the following information: (1) study identification, (2) study methodology, (3) research participants, (4) effect sizes and dependent variables, and, for intervention meta-analyses, (5) the characteristics of the interventions (Wilson, 2009).

Study Identification

Study identification involves recording the identifying characteristics of the studies in the meta-analysis. This information typically includes the title, date, and author(s) of the study; the source of the study (e.g., journal article, conference paper, doctoral dissertation, technical report, book chapter); retrieval source (e.g., computer search, reference list); and study setting or region.

Study Methodology

While methodological characteristics of studies may not be of substantive interest to practitioners or policy makers, they are a critical part of coding because the methods used to conduct research studies can influence their outcomes. For example, studies in which subjects are randomly assigned to treatment and comparison groups tend to have different effects than studies in which subjects are not randomly assigned (Lipsey & Wilson, 2001). Coding details of the study methods allows the meta-analyst to identify influential method characteristics and control for them in any statistical analysis. To illustrate, imagine a meta-analysis of therapeutic interventions for anxiety where, in the primary studies, cognitive restructuring programs tended to use more randomized experiments, while psychoanalytic programs tended to be evaluated more often with quasi-experimental designs. If the average effect size for the cognitive restructuring programs turns out to be larger than the average effect size for the psychoanalytic treatments, we cannot be sure whether cognitive restructuring programs are more effective than psychoanalysis or whether randomized experiments might tend to result in larger program effects regardless of the type of treatment. During analysis of results, having information about research design and other methodological characteristics can be immensely useful for describing the overall methodological quality of the candidate studies, separating the results for high- and low-quality studies, and using statistical methods to control for confounding methodological differences between studies.

Many different aspects of method and procedure can be coded in a meta-analysis or systematic review. Some meta-analyses may refer to methodological quality checklists that are available (Chalmers et al., 1981; Valentine, 2009); others may frame the methodological coding around the validity of the findings or the risk of bias present in the candidate studies (Higgins & Altman, 2008). Still others may perform an objective coding of the methodological characteristics of interest and examine their influence on the effect sizes later. Some common methodological characteristics included in the various meta-analysis coding schemes include

method of assignment to research conditions (e.g., random assignment, cluster randomization, matching), nature of the comparison group (e.g., received no treatment, placebo, alternate treatment), study attrition, study blinding, pretreatment equivalence of groups, outcome measures used to assess treatment effects (e.g., norm referenced, criterion referenced, rating scale), timing of follow-up outcome measures, and reliability of outcome measures.

Research Participants

Coding candidate studies for participant characteristics allows the meta-analyst to determine whether included studies have similar target populations and subsequently examine whether study findings are associated with those participant characteristics. This information is important for assessing the appropriateness of different interventions for use with different types of participants and for understanding how the etiology of social problems might differ across participant subgroups. A great variety of participant characteristics may be coded, depending on the research questions and the types of participant information commonly reported in the literature being reviewed. In general, basic demographic information including gender, race/ethnicity, age or grade, and socioeconomic status is important. Risk status, diagnostic status, severity of problem behavior, education level, previous health and mental health histories, and a range of other personal characteristics might also be relevant. The information coded about research participants should be informed by the research questions and theory or empirical research that identifies participant characteristics that may be associated with the outcomes of interest.

Intervention Characteristics

For meta-analyses of intervention effects, coding characteristics of the intervention programs is of primary importance, both for understanding the quality of the treatment's implementation and for determining the overall effectiveness of the treatment program(s). Examples of treatment program characteristics used to examine variation in intervention effects include length of treatment program, number of treatment sessions in program, length of each treatment session, type of treatment program, individual or group treatment sessions, fidelity of treatment (i.e., treatment implemented as described), and type and training of treatment administration personnel.

Effect Sizes

Finally, perhaps the most critical pieces of information extracted during the coding phase of a meta-analysis are those related to the effect sizes (i.e., the actual statistical results of the study). The effect sizes most commonly used in prevention

science fall into three major families: those that index differences between groups on continuous measures (e.g., the standardized mean difference or Cohen's *d*), those that index relationships between two continuous measures (e.g., the correlation coefficient), and those that index differences between groups on frequency or incidence (e.g., odds ratio or risk ratio).

In any coding scheme, each effect size that is coded will be accompanied by a set of codes that provide detailed information about its distinctive source and nature. Such information might include the data used in the computation, sample size on which the effect size is based, amount of attrition in the sample, manipulations used (e.g., to derive it from other statistics), and other such items that can be used in statistical analysis to examine methodological or procedural matters that may systematically influence effect size.

Dependent Variables

The coded information for each effect size also typically includes information about the variables involved in the index. For instance, this coding could identify the construct represented in the index, the nature of the behavior at issue, source of the information, and relevant features of the operationalization. It is this coding that forms the basis for grouping the effect sizes that are combined in a given analysis, that is, those that are treated as representing the "same" construct for that particular analysis.

Potential Bias at the Coding Stage

As with the eligibility criteria, construct ambiguity can introduce bias at the coding stage of a meta-analysis. The validity of conclusions from a meta-analysis can also be threatened by issues of reliability that arise during the coding of eligible study reports. When multiple coders extract the same type of information from multiple studies, there is always the risk of unreliability. Therefore, it is important for meta-analysts to conduct extensive training sessions with coders and to assess the reliability of coding. Ideally, all study reports would be coded by two independent coders, and all discrepancies would be resolved through discussion and further training. In practice, however, it may not be feasible (due to budgetary reasons or otherwise) to have all studies double coded; in this case, it is extremely important for inter-coder agreement to be established early on during the project and continuously monitored with subsets of studies. Validity of conclusions can be threatened when multiple coders are not reliable with one another and when individual coders are not reliable with themselves (what has been called "coder drift"). Coder drift can be the result of coder fatigue or a change in understanding of constructs over time. Therefore, it is important to conduct, at minimum, continuous monitoring of subsets of coded studies to quickly assess and remedy any reliability problems with the coding.

Data Analysis and Interpretation

As with any form of research, the final stage of the project comes when the data are analyzed and interpreted and conclusions are drawn about the body of research under study. Analysis of meta-analytic data has several statistical quirks but proceeds much like analysis of data from primary studies. Several good texts are available, and a variety of software packages and macros have been developed to handle the specific issues associated with analyzing effect sizes (Borenstein et al., 2009; Borenstein, Hedges, Higgins, & Rothstein, 2010; Lipsey & Wilson, 2001; Sterne, 2009), so we won't spend time here reviewing the analytic techniques in detail. Rather, the remainder of this section provides a basic overview of meta-analysis methods and some applications to prevention science.

In addition to summarizing the basic study characteristics of the literature reviewed, a typical prevention science meta-analysis would include the following components: (1) the average effect size and effect size distribution for each outcome of interest and an examination of the heterogeneity in the effect size distributions, (2) subgroup or moderator analysis in which the variability present in the effect size distribution is systematically analyzed to identify study characteristics that are associated with larger or smaller effect sizes, and (3) publication bias analysis and other sensitivity analyses to assess the validity of conclusions drawn. We briefly review each of these in turn.

Average Effect Sizes and Heterogeneity

Most meta-analyses will present an average effect size value synthesized from the individual effect sizes extracted from the primary studies included in the review. When calculating the average effect size, each effect size is typically weighted by the inverse of its sampling variance, so that effect sizes measured with greater precision are given greater weight because they provide better estimates of the underlying population parameter(s) of interest. Meta-analysts will typically provide estimates of average effect sizes and their distribution for each outcome of interest to make broad statements about the average effect in the population.

For instance, in their meta-analysis examining the effects of social information-processing programs on students' aggressive and disruptive behavior, Wilson and Lipsey (2006) reported an overall average standardized mean difference effect size of 0.21 based on 73 studies and concluded that the intervention had a small but statistically significant effect on reducing youths' aggressive and disruptive behavior. By calculating the average effect size across all included studies, meta-analysts can thus answer the simple question of "did the interventions work?" (or for other types of research questions, "is there an association between two variables?" or "are there differences between these groups?").

In many cases, the meta-analyst may be interested, not only in the average effect, but also in the variability of those effects across different types of studies,

participant samples, and so forth. Therefore, most meta-analyses also present statistics that summarize the amount of variability between studies, test whether any observed heterogeneity may be due to chance, and summarize the proportion of observed heterogeneity that can be considered true heterogeneity rather than statistical noise (i.e., estimates of τ^2, Q or χ^2, and I^2, respectively; see Borenstein et al., 2009). When heterogeneity statistics indicate that substantial heterogeneity is present, some meta-analysts may decide that the studies are, in fact, too heterogeneous to calculate an average effect or do any meaningful statistical synthesis. Although this decision may be justified in some situations, many prevention scientists may actually be interested in this heterogeneity and, therefore, choose to empirically examine it through the use of subgroup or moderator analysis. The Wilson and Lipsey (2006) review, for instance, found substantial heterogeneity in their effect size distribution and, thus, proceeded to statistically examine a variety of factors that may have contributed to that heterogeneity.

Moderator/Subgroup Analysis

In meta-analysis, moderator analysis refers to statistical analyses that examine whether the coded study characteristics for each study are associated with the effect sizes from those studies, that is, whether coded variables can *explain* some or all of the observed heterogeneity in the effect sizes (see Lipsey & Wilson, 2001, for more technical detail). This type of analysis is called "moderator" analysis in that it examines whether a certain coded variable or variables (x) are associated with the direction or magnitude of the effect size (y), when the effect size is defined as an index of the association between two variables (i.e., the association between a treatment variable and an outcome variable in an intervention meta-analysis, the association between a predictor and an outcome in a epidemiological meta-analysis, or the association between group membership and an outcome variable in a group differences meta-analysis). Thus, the covariate x is framed as a moderator of the relationship between the two variables encapsulated in the effect size y. Moderator analysis is conducted using analogs to ANOVA and linear regression that are modified for use with meta-analytic data. The choice between the ANOVA and regression frameworks depends on the measurement level of the covariate(s) of interest. Typically, "subgroup" analysis refers to a moderator analysis of categorical covariates in the ANOVA analog framework.

Moderator analysis thus allows meta-analysts to examine myriad factors that may be associated with the study findings. Questions that may be particularly relevant to prevention scientists include examining the conditions under which and for whom certain interventions may be most effective, certain risk factors may be most important, or differences between groups may be largest. For instance, the Wilson and Lipsey (2006) meta-analysis found that studies with primarily low socioeconomic status participants, those with more frequent treatment contact with participants, those in which interventions were delivered for research and demonstration purposes, and those with no obvious implementation difficulties produced

the largest intervention effects. Opportunities for moderator analysis in a meta-analysis are limited only by the size of the literature under review and the characteristics of the study variables coded for each study. That said, however, it is important for meta-analysts to identify moderators of interest a priori—not only so that the moderators can be coded during the data collection phase, but also to minimize any data dredging at the analysis phase that might capitalize on chance.

Returning to the three general types of research questions outlined above, meta-analyses that focus on questions of etiology, epidemiology, and the development of social problems present several interesting analysis opportunities. First, these correlational meta-analyses produce a quantitative summary of the strength of relationships among the variables of interest. That information can be fed into analyses that examine the differential predictive strength of different risk or protective factors or variation in the predictive strength of risk or protective factors for individuals or groups with different characteristics. This information can be used to identify both the target behaviors for intervention as well as the best individuals to target for intervention services. In a similar fashion, group differences meta-analyses lend themselves to questions about diagnostic groups that might be particularly amenable to treatment or at particular risk for later problems or comorbidities; they can also identify intervention targets that may differ across different subgroups of the population. Finally, intervention meta-analyses provide a variety of analysis opportunities for producing policy-relevant results. Questions of what works best under what conditions, in what types of settings, and for what types of individuals are important here and are most defensible when meta-analysts do a careful assessment of the risk of bias of the included studies and are also mindful of the influence that study methods can have on research findings.

Publication Bias and Sensitivity Analysis

As previously mentioned, researchers must acknowledge the possibility of publication bias and how it may affect the results of a meta-analysis. There are several exploratory statistical procedures that meta-analysts can use to examine the possibility of publication bias (see Rothstein et al., 2005). The most commonly reported procedures include visual inspections of funnel plots or regression-based tests for effect sizes based on continuous outcome data (Egger, Davey Smith, Schneider, & Minder, 1997) or dichotomous outcome data (Harbord, Egger, & Sterne, 2006; Peters, Sutton, Jones, Abrams, & Rushton, 2006; Rucker, Schwarzer, & Carpenter, 2008). Other methods that have been used to assess publication bias in the past, such as the rank correlation test and variations on the fail-safe *N*, are no longer recommended for use given their known limitations (Becker, 2005; Sterne, Gavaghan, & Egger, 2000). Statistical development of publication bias analysis methods is constantly evolving, however, and there are currently no consistently agreed-on standards in the field. This, along with the known limitations of currently used methods (e.g., low power), means that there is no simple solution for detecting publication bias in a meta-analysis. Nonetheless, meta-analysts must be sensitive to

the possibility of publication bias and its potential effect on the conclusions that can be drawn from the results.

In addition to conducting exploratory analysis to assess the possibility of publication bias in a meta-analysis, it is also common practice for meta-analysts to conduct sensitivity analyses. As with any analysis of data in a primary study, the meta-analyst makes decisions during the data collection and analysis phases that could conceivably influence study results and conclusions. Therefore, it is important to conduct sensitivity analyses that explore whether those decisions had an appreciable impact on the meta-analysis findings. For instance, sensitivity analyses may explore the impact of (1) only including randomized studies of intervention effectiveness, (2) only including studies published in English, (3) the use of any statistical adjustments (4) how outlier cases were handled, (5) potential confounding among moderators, (6) how missing data were handled, and so on.

Potential Biases at the Data Analysis Stage

Several threats to validity can occur at the analysis stage of a meta-analysis. A comprehensive understanding of these threats presumes a working knowledge of meta-analytic methods and standards, of course, but we will review some of the more common issues here (we again refer interested readers to comprehensive books on meta-analysis methods for a further understanding of the issues at hand, e.g., Borenstein et al., 2009; Cooper, 2010; Cooper et al., 2009; Lipsey & Wilson, 2001; Littell, Corcoran, & Pillai, 2008).

Effect Size Approximations

The statistical assumptions and methods employed in a meta-analysis can influence the validity of the conclusions. First, there is the issue of computing effect size estimates—the currency and primary outcome in a meta-analysis. Most meta-analysts can attest to the frustrating reality of discovering that the information needed to calculate an effect size statistic or its corresponding standard error (e.g., sample sizes, standard deviations) has been omitted from a primary study. Meta-analysts should be transparent about how they deal with such missing data and the extent to which any effect size estimates were approximated from partially reported data. For instance, one common effect size index, the standardized mean difference (or Cohen's d), is easily computed when study authors report sample sizes, means, and standard deviations. Algebraically equivalent formulas are available for calculating d based on other statistics, such as t-tests and F-tests, but when such information is not available, meta-analysts may choose to estimate d using formulas for algebraic approximations based on other pieces of information (e.g., statistics from a two-factor repeated measures ANOVA). The accuracy of these approximations may bias effect size estimates. However, if researchers code information regarding the estimation method or level of approximation needed to

calculate the effect size, it is possible to examine these variables as moderators or statistically control for them in the final analyses.

Study Quality

Another important issue at all stages of meta-analysis, but particularly so at the data analysis phase, is the "quality" of the studies included in the analysis and how that may bias the findings. Indeed, the results of a meta-analysis are entirely dependent on the primary studies of which it is comprised. Many meta-analyses have been legitimately criticized for synthesizing results from low-quality studies. This "garbage-in-garbage-out" criticism of meta-analysis emphasizes the need for a careful consideration of the quality of those studies included and analyzed. The difficulty lies in defining quality, however, and there is no simple definition of quality. There are many tools available that can be used to assess the "risk of bias" that may result from including lower-quality studies in a meta-analysis (Valentine, 2009). The Cochrane Collaboration systematic reviews, for instance, include risk of bias tables that assess the possibility of bias due to the quality (or lack thereof) of the included primary studies (Higgins & Altman, 2008). Most risk of bias instruments are geared toward meta-analyses of randomized controlled trials, however, and are less applicable to etiological/epidemiological or group differences meta-analyses. Therefore, it is also common for meta-analysts to assess study quality in a post hoc, empirical fashion during the analysis stage. Namely, by coding information on study quality such as measurement validity and reliability, implementation fidelity, attrition, blinding, and so forth, the meta-analyst can conduct subgroup or moderator analyses to examine whether those quality variables are associated with the effect size or explain some of the heterogeneity across studies that might be associated with the quality indicators.

Dependent Effect Sizes

Meta-analysis methods assume the independence of effect size estimates (i.e., that any given analysis only includes one effect size estimate per study). However, many primary studies report enough information to calculate multiple effect size estimates. For example, a researcher may propose to conduct a meta-analysis summarizing the effects of a life skills prevention program on high school students' alcohol use. It is plausible that several eligible studies may include multiple measures of "alcohol use." A study may include information on the number of days a student drank any alcohol in the past 30 days, the number of days a student drank any alcohol in the past 90 days, the amount of alcohol consumed on a given occasion, and so forth. Even if the meta-analyst decided these were all eligible outcomes representing the same underlying alcohol use construct of interest, most meta-analysis methods assume the use of only one effect size estimate per study in any given analysis.

To avoid statistical dependencies, most meta-analysts use one of several techniques: (1) they create one average effect size per study, thereby losing the ability to account for the distinct characteristics of the different outcomes; (2) they choose one effect size per study based on some decisional criteria, thereby throwing away information contained in the other effect sizes; or (3) they conduct several separate meta-analyses split by some characteristics of the outcomes or effect sizes, thereby preventing comparison of common moderators. Historically, the only statistically defensible alternative if the meta-analyst wished to include all effect sizes in the same analysis has been to model the dependencies among effect size estimates drawn from the same study (Gleser & Olkin, 2008). This is rarely feasible because it requires information about the correlations among the outcome variables that are virtually never reported by primary study authors. There is a newly developed technique (Hedges, Tipton, & Johnson, 2010) that estimates robust standard errors that can adjust for the lack of statistical independence so that all relevant outcome variables measuring the same outcome construct can be used in the same analysis. This is a new technique, however, and is not yet widely used. Unfortunately, many meta-analysts have altogether ignored the issue of dependent effect size estimates, incorrectly including multiple effect size estimates from the same study in an analysis, with no adjustment for the fact that doing so violates the assumptions of meta-analytic techniques. Doing so can yield incorrect results. In some cases, this underestimates standard errors and increases the possibility of finding a significant effect; in other cases, standard errors may be overestimated. It is therefore important for consumers of meta-analyses to understand whether a meta-analysis has correctly analyzed dependent effect size estimates.

Summary

In conjunction with primary studies of etiology, epidemiology, intervention effectiveness, and group differences, meta-analysis is an important analytic method for use in the field of prevention science. By allowing researchers to systematically summarize the empirical literature in a given area, meta-analysis can be a powerful tool for informing the science of prevention. Meta-analyses that study the etiology of social problems, the effectiveness of interventions, and the differences between groups can identify high-risk groups or those in need of services, and they can aid in the development and implementation of primary, secondary, and tertiary prevention programs, most especially in the contexts under which they may have the largest impact.

References

Bálint, S., Czobor, P., Komlósi, S., Mészáros, Á., Simon, V., & Bitter, I. (2009). Attention deficit and hyperactivity disorder (ADHD): Gender- and age-related differences in neurocognition. *Psychological Medicine, 39*, 1337–1345.

Battin-Pearson, S., Newcomb, M. D., Abbott, R. D., Hill, K. G., Catalano, R. F., & Hawkins, J. D. (2000). Predictors of early high school dropout: A test of five theories. *Journal of Educational Psychology, 92*, 568–582.

Becker, B. J. (2005). Failsafe N or file-drawer number. In H. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 111–125). West Sussex, UK: Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*, 97–111.

Catalano, R. F., Hawkins, J. D., Berglund, M. L., Pollard, J. A., & Arthur, M. W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*(Suppl 6), 230–239.

Chalmers, T. C. (2001). Using systematic reviews and registers of ongoing trials for scientific and ethical trial design, monitoring, and reporting. In M. Egger, D. Smith, & D. G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (2nd ed., pp. 429–443). London, UK: BMJ Books.

Chalmers, T. C., Smith, H., Jr., Blackburn, B., Silverman, B., Schroeder, B., Reitman, D., et al. (1981). A method for assessing the quality of a randomized control trial. *Controlled Clinical Trials, 2*, 31–49.

Cooper, H. (1984). *The integrative research review: A systematic review*. Newbury Park, CA: Sage.

Cooper, H. (2010). *Research synthesis and meta-analysis* (Applied social research methods series 4th ed., Vol. 2). Thousand Oaks, CA: Sage.

Cooper, H., DeNeve, K., & Charlton, K. (1997). Finding the missing science: The fate of studies submitted for review by a human subjects committee. *Psychological Methods, 2*, 447–452.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage.

Dickersin, K. (1997). How important is publication bias? A synthesis of available data. *AIDS Education and Prevention, 9*, 15–21.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634.

Egger, M., Zellweger-Zahner, T., Schneider, M., Junker, C., Lengeler, C., & Antes, G. (1997). Language bias in randomized controlled trials published in English and German. *Lancet, 350*, 326–329.

Ekeland, E., Heian, F., Hagen, K. B., Abbott, J., & Nordheim, L. (2005). Exercise to improve self-esteem in children and young people. *Campbell Systematic Reviews, 4*.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3–8.

Gleser, L. J., & Olkin, I. (2008). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 357–376). New York: Russell Sage.

Gotzsche, P. C. (1987). Reference bias in reports of drug trials. *British Medical Journal, 195*, 654–656.

Harbord, R. M., Egger, M., & Sterne, J. A. (2006). A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine, 25*, 3443–3457.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*, 39–65.

Higgins, J. P. T., & Altman, D. G. (2008). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions (chapter 8* (pp. 187–241). Chichester, UK: Wiley.

Higgins J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions* (Version 5.1.0, updated March 2011). The Cochrane Collaboration. Retrieved from http://www.cochrane-handbook.org

Hopewell, S., & Clarke, M. (2001). Methodologists and their methods: Do methodologists write up their conference presentations or is it just 15 minutes of fame? *International Journal of Technology Assessment in Health Care, 17*(4), 601–603.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies*. Beverly Hills, CA: Sage.

Kellam, S. G., Koretz, D., & Mościcki, E. K. (1999). Core elements of developmental epidemiologically based prevention research. *American Journal of Community Psychology, 24*, 463–482.

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

Lindberg, S. M., Hyde, J. S., Linn, M. C., & Petersen, J. L. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*, 1123–1135.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*, 1181–1209.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Littell, J. A., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*. New York: Oxford University Press.

Matt, G. E., & Cook, T. D. (2009). Threats to the validity of research syntheses. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (2nd ed., pp. 537–560). New York: Russell Sage Foundation.

Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2006). Comparison of two methods to detect publication bias in meta-analysis. *Journal of the American Medical Association, 295*, 676–680.

Petrosino, A., Petrosino, C. T., & Buehler, J. (2004). "Scared Straight" and other juvenile awareness programs for preventing juvenile delinquency. *Campbell Systematic Reviews, 2*.

Pittler, M. H., Abbot, N. C., Harkness, E. F., & Ernst, E. (2000). Location bias in controlled clinical trials: A survey of three medical journals. *Journal of Clinical Epidemiology, 2000*(53), 485–489.

Rennie, D. (1997). Thyroid storm. *Journal of the American Medical Association, 277*, 1238–1243.

Reyes, M. M., Panza, K. E., Martin, A., & Bloch, M. H. (2011). Time-lag bias in trials of pediatric antidepressants: A systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(1), 63–72.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638–641.

Rothstein, H. R., & Hopewell, S. (2008). Grey literature. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 103–125). New York: Russell Sage.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias analysis: Prevention, assessment, and adjustments*. West Sussex, UK: Wiley.

Rucker, G., Schwarzer, G., & Carpenter, J. (2008). Arcsine test for publication bias in meta-analyses with binary outcomes. *Statistics in Medicine, 27*, 746–763.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist, 47*, 1173–1181.

Simes, R. J. (1986). Publication bias: The case for an international registry of clinical trials. *Journal of Clinical Oncology, 4*, 1529–1541.

Sterne, J. A. C. (Ed.). (2009). *Meta-analysis in Stata: An updated collection from the Stata Journal*. College Station, TX: Stata Press.

Sterne, J. A. C., Egger, M., & Moher, D. (2008). Addressing reporting biases. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions (chapter 10)*. Chichester, UK: Wiley.

Sterne, J. A. C., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology, 53*, 1119–1129.

Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2011). Risk factors and crime. In F. Cullen & P. Wilcox (Eds.), *The Oxford handbook of criminological theory*. New York: Oxford University Press.

Tannock, I. F. (1996). False-positive results in clinical trials: Multiple significance tests and the problems of unreported comparisons. *Journal of the National Cancer Institute, 88*, 206–207.

Valentine, J. C. (2009). Judging the quality of primary research. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 129–146). New York: Russell Sage.

Wilson, D. B. (2009). Systematic coding. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 159–176). New York: Russell Sage.

Wilson, S. J., & Lipsey, M. W. (2006). The effects of school-based social information processing interventions on aggressive behavior, part I: Universal programs. *Campbell Systematic Reviews*, **5**.

Chapter 20
Mixed Methods Research Design for Prevention Science: Methods, Critiques, and Recommendations

Felipe González Castro, Osvaldo F. Morera, Joshua G. Kellison, and Katherine M. Aguirre

Background on Mixed Methods Research for Prevention Science

This chapter examines the complex and challenging considerations involved in the design of rigorous mixed methods research studies. In this effort, research investigators must understand and address certain competing or conflicting qualitative and quantitative traditions and imperatives for conducting a scientifically rigorous, yet culturally sensitive mixed methods research study. The aim is to design a well-crafted mixed methods research study that incorporates the best elements of qualitative and quantitative data forms for a study that generates more informative prevention science research.

The "paradigm wars." Historically, the diverse and at times conflicting schools of thought and traditions that have created the contemporary field of mixed methods research provide a rich background for informing the design of a well-crafted mixed methods research study. These diverse qualitative–quantitative perspectives introduce distinct epistemological assumptions and epistemological tensions that have been described as the "paradigm wars." These dynamic tensions have been prominent as expressed within three historical recent eras: (a) the era of conflicts between the objective, analytic scientific perspective branded as "positivism" and the subjective, interpretive perspective labeled as "constructivism"; (b) the era of "competing positivist, constructivist, and critical theory paradigms"; and (c) the contemporary era of conflicts between "evidence-based methodologies" and the "interpretive and critical theory schools" (Denzin, 2012).

Despite these historical paradigm clashes that framed the positivist–constructivist dichotomy, an integrative perspective as offered by contemporary multicultural psychology recognizes considerable compatibility in the scientific and

F.G. Castro (✉)
Department of Psychology, University of Texas, 500 W. University, El Paso, TX 79968, USA
e-mail: fcastro4@utep.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_20,

constructivist perspectives as evident within contemporary research studies that examine the "world views" of diverse people of color (Sue & Sue, 1999). Regarding these world views, it has been noted that "the complex nature of the social world requires a more fluid understanding and application of the relationship between philosophical paradigms (assumptions about the social world and nature of knowledge), methodology (the logic of inquiry), and methods (techniques and data collection)" (Sharp et al., 2012, p. 36).

That said, a purely philosophical analysis of these competing paradigms, while instructive, nonetheless lacks praxis (i.e., useful application). In other words, such philosophical analyses do not specify the procedures for "mixing," "blending," "combining," or otherwise "integrating" evidence from these distinct data sources to generate credible results and for drawing clear and convincing conclusions that inform real-world applications for solving pressing human and social problems. From a practical perspective, and as applied to prevention science research, a significant and overarching research aim is to apply evidence-based knowledge about "what works and how it works" and, whether based on qualitative or quantitative data, to inform the design of prevention interventions that effectively solve abiding social and public health problems.

A pragmatic approach. From this contextual backdrop, our approach to the conduct of mixed methods research as applied to prevention science is best regarded as *pragmatism*. In taking this position, we readily acknowledge the existence of other paradigms and traditions and their differing approaches for conducting mixed methods research and research in general. Pragmatism in research emphasizes action over philosophizing and endorses practical theory or *praxis* as an approach for informing effective practice (Sharp et al., 2012). We also recognize that pragmatism as an approach will not be attractive to certain scholars from certain mixed methods communities, particularly those who are philosophically opposed to the so-called positivistic, scientific approach. Beyond this, within the field of prevention science, the aim is to advance beyond a particular paradigm, to design rigorous research studies that generate robust results, advancing beyond polemics to the application of efficacious evidence-based interventions that produce genuine and sustainable behavioral, social, and/or policy changes. We also note that regardless of the paradigm used, a poorly crafted research study will fail to generate convincing arguments, informative results, and scientifically sound conclusions.

The approach. Accordingly, within this chapter, our aim and approach involve examining recent studies from a scientific perspective and in accord with prevention science standards of evidence (Flay et al., 2005). Our view is that a strong research design emerges from a clear theoretical or conceptual framework and from a clear research rationale that includes an explicitly described research plan. Moreover, from a rigorous and integrative mixed methods perspective (Castro, Kellison, Boyd, & Kopak, 2010), a well-specified research plan for mixing methods *systematically* is most important and perhaps indispensable to full qualitative–quantitative data integration (Fielding, 2012). By extension, a major goal in the design of well-crafted mixed methods research studies involves maximizing rigor

while allowing some methodological flexibility that incorporates the respective strengths of quantitative and qualitative data forms and minimizes their weaknesses. Optimizing the contributions of qualitative and quantitative data forms often involves adjustments and tradeoffs, although also with internal consistency in design, results, and interpretations, one that generates clear conclusions that are closely grounded in the empirical evidence generated from that research study (Fielding, 2012; Tashakkori & Teddlie, 2003). The quality of that evidence is thus a product of a "thoughtfulness about data integration that inspires confidence in findings so that the research makes a difference in the real world" (Fielding, 2012, p. 134).

What Is Mixed Methods Research?

Defining mixed methods research. Mixed methods research has emerged recently as a set of research methodologies and as a field of scholarly inquiry (Teddlie & Tashakkori, 2003). This field incorporates multiple epistemological and methodological approaches to combine information obtained from quantitative (numeric) and qualitative (text narrative) data. Both data forms constitute "evidence" (Castro & Coe, 2007) that can vary in research quality and that in combination can generate research results that are not obtainable from the use of either data form alone (Tashakkori & Teddlie, 2003). These data can be gathered *concurrently* (within a single stage or phase) or *sequentially* (in two temporal stages or phases). Given the diverse perspectives that have created the mixed methods field, no single definition aptly captures the complexity and breadth of the diverse research studies that constitute mixed methods research. Nonetheless, based on the perspectives of major leaders in the mixed methods research field, an emerging definition is that:

> Mixed methods research is an intellectual and practical synthesis based on qualitative and quantitative research; it is the third methodological or research paradigm (along with qualitative and quantitative research). It recognizes the importance of traditional quantitative and qualitative research but also offers a powerful third paradigm choice that often will provide the most informative, complete, balanced, and useful research results. (Johnson, Onwuegbuzie, & Turner, 2007, p. 129)

Accordingly, a benefit from mixed methods research is *complementarity*, the gains obtained from the "complementary strengths" offered by two differing data forms to produce results having greater explanatory strength (Johnson et al., 2007). In principle, a well-designed mixed methods research study offers the respective strengths of quantitative and of qualitative methods while minimizing their weaknesses. However, in practice, many contemporary studies labeled as "mixed methods" lack research rigor and trustworthiness (Fielding, 2012; Krefting, 1991). This major limitation also includes a lack of qualitative–quantitative data integration, which often has diminished the quality and significance of the research results generated.

Strengths and limitations of qualitative and quantitative methods. As a strength, quantitative multivariate model analyses can generate *confirmatory* results, that is, outcomes high in internal validity and allowing precise conclusions or answers to one or more specific research questions. A relative weakness of the *quantitative* approach is that results may lack depth of information or analysis, thus offering precise but shallow explanations. By contrast, a strength of *qualitative* methodologies is that they can generate rich and detailed descriptions about complex human thoughts, emotions, and experiences. Conversely, a relative weakness of *qualitative* approaches is that they often disallow formal hypothesis testing and may generate difficult to interpret research results (Dreher, 1994). In addition, textual evidence can be difficult to integrate across cases. Thus, the qualitative analysis of complex or divergent research evidence can yield conflicting results, although such disparate or divergent results may still be analyzed in a manner that uncovers hidden information and can produce more nuanced interpretations (Hesse-Biber, 2012).

Maximizing internal and external validity. From the quantitative research tradition, a study high in *internal validity* can generate accurate conclusions (i.e., clarity of causal inference) (Flay et al., 2005). Study results are thus attributable with high confidence to the effect of interest, based on ruling out competing threats to causal inference (Flay et al., 2005; Patten, 2012). By contrast, the selection of a biased or nonrepresentative sample compromises *external validity*, limiting or disallowing the generalization of study results to a larger population or in a manner prescribed a priori by the purpose of the research study (Sharp et al., 2012). Within this context, we acknowledge the distinct and at times conflicting standards of evidence that have emerged historically from the noted qualitative and quantitative research traditions. As a hybrid of both research traditions, mixed methods research is thus affected by a dynamic tension resulting from efforts to design a research study that considers the views of these distinct traditions. Nonetheless, from their own perspectives, both traditions emphasize the importance of *research rigor*, as defined in different ways. From the qualitative tradition, rigorous research should exhibit *trustworthiness*, as indicated by the criteria of (a) truth value, (b) applicability, (c) consistency, and (d) neutrality (Krefting, 1991; Milinki, 1999), as assessed respectively by the research criteria of (a) credibility, (b) transferability, (c) dependability, and (d) confirmability. It is noteworthy that these four criteria stand in parallel to respective criteria for rigor in *quantitative* research: (a) internal validity, (b) external validity, (c) reliability, and (d) objectivity. These parallels in qualitative and quantitative criteria for research rigor form the foundation for an integrative application of these criteria to establish rigor in the design of well-crafted and genuine mixed methods research studies.

In principle, mixed methods research designs should be capable of attaining both high internal and high external validity. By contrast, in the past, mixed methods research designs have often been methodologically weak, exhibiting both low internal and external validity, in addition to other research design flaws or limitations. Despite the challenges of developing well-designed mixed methods research studies, such studies can offer the following benefits: (a) *triangulation,* the use of two or more methods for examining a research problem; (b) *complementarity,* the

use of two data forms to obtain a deeper understanding of a research problem; (c) *development,* a synergistic effect in which the results of one method can inform the other; (d) *initiation*, whereby findings can raise questions or reveal contradictions that require clarification, thus prompting new insights; and (e) *expansion,* a broader and more profound approach to inquiry (Hesse-Biber, 2010).

Furthermore, mixed methods approaches are useful for "giving voice" to members of special populations (e.g., injection drug users, racial/ethnic minorities, men who have sex with men [MSMs], women victimized by domestic violence, lesbian and gay families) (Hesse-Biber, 2012; Lassiter, Dew, Newton, Hays, & Yarbrough, 2006; Mertens, 2003). Mixed methods approaches are also useful for studying the richness and complexity of "culture" within its natural environmental contexts (Gonzalez-Guarda, Florom-Smith, & Thomas, 2011), especially as complex cultural effects are not fully captured or understood by the sole use of numeric-scaled data, even when using multilevel systemic and/or structural equation models.

These benefits of mixed methods research studies do not diminish the value and importance of established multivariate data analytic methods (Cohen, Cohen, West, & Aiken, 2003; Hayes, 2013; Preacher, Zyphur, & Zhang, 2010). Moreover, mixed methods research, when rigorously designed and implemented, offers an innovative and expanded methodology that can complement the use of multivariate analytic methods. When applied with special populations, mixed methods approaches can generate novel results that are not easily captured by the sole use of conventional multivariate methods or conventional qualitative methods (Yoshikawa, Weisner, Kalil, & Way, 2008).

Informing complex systems models and processes. Well-designed mixed methods research may provide exploratory evidence to inform studies of complex systems and processes in the field of systems science. For example, *systems dynamic modeling* examines complex systems and their dynamic processes of change (Homer & Hirsch, 2006) in a manner not attainable with conventional statistical modeling (Diez Roux, 2011). For example, using a 2-stage study, during stage 1, mixed methods approaches can generate data to explore complex processes and interactions that may occur within an extended family system, as this can inform the development of a more specific dynamic model simulation that would be tested in stage 2. Similarly, *agent-based modeling* is another systems science simulation methodology that examines the behaviors of a system of agents and the interactions between them (Tang & Bennett, 2010). Examples include patterns of learning, reactions to novel situations, and adaptations to "emergent phenomena" to examine how dynamic changes in behavior can affect certain outcomes (Siebers, Macal, Garnett, Buxton, & Pidd, 2010; Tang & Bennett, 2010). Here also a well-designed exploratory mixed methods study conducted in stage 1 can explore complex environmental and person effects and their interactions to inform the design of an agent-based model simulation that would be conducted in stage 2.

The Emergence of Mixed Methods Research Conceptualization of the Implementation Process

Within the past decade, mixed methods methodologies have emerged as viable and more broadly accepted research approaches. Indicators of this growth and acceptance include the first major book on mixed methods approaches, *Handbook of Mixed Methods in Social and Behavioral Research* (Tashakkori & Teddlie, 2003, 2010), and the establishment of a journal dedicated to research in the mixed methods field, *Journal of Mixed Methods Research*, initiated in 2007. In addition, Division 5 (Evaluation, Measurement, and Statistics) of the American Psychological Association recently voted to include the Society of Qualitative Inquiry in Psychology (SQIP) as a part of the Division. A stated reason for the inclusion of SQIP was the potential to strengthen mixed methods research (Little & Freeman, 2011).

As another milestone, the National Institutes of Health (NIH) and the Office of Behavioral and Social Science Research (OBSSR) have emphasized the importance of mixed methods approaches for understanding technology transfer within the emerging field of dissemination and implementation research (Brown, Colditz, & Proctor, 2012; Landsverk et al., 2012). Also, OBSSR has sponsored the development of a working group report titled *Best Practices for Mixed Methods Research in the Health Sciences* (Creswell, Klassen, Plano Clark, & Smith, 2011).

Relevance of Mixed Methods to Prevention Science

The approach. Scientific research, as endorsed by NIH and as practiced in the field of prevention science, has been guided by five research criteria: significance, innovation, approach, investigators, and the research environment (Center for Scientific Review, 2013). Within this context, the design and implementation of scientifically meritorious mixed methods research studies involve the challenge of integrating core features of the qualitative and quantitative traditions, which at times introduces conflicting imperatives (Creswell et al., 2011; Flay et al., 2005). Nonetheless, well-designed mixed methods research studies can extend conventional research designs to yield research results that better inform prevention science research in its three major areas: (a) epidemiology, (b) intervention development, and (c) research methodology (Biglan et al., 2011).

A few caveats. Despite the potential benefits of mixed methods research, one caveat is that "the work is just beginning with regard to mixing the study of 'words' and 'numbers' in scientific research" (Yoshikawa et al., 2008, p. 352). It has also been noted that until recently, "the fundamental issue of the degree to which mixed methods researchers genuinely integrate their findings has not been addressed to a significant extent" (Bryman, 2007, p. 8). However, as these caveats can be addressed effectively in the design and implementation of rigorous mixed methods

research studies, such studies can complement the use of conventional multivariate methodologies to better inform prevention science research.

Mixed Methods Research Designs for Rigor in Mixed Methods Studies

Overview. One "best practice" recommendation from the recent OBSSR working group report is that research investigators should specify the particular mixed methods research design that they will use in a proposed mixed methods research study (Creswell et al., 2011). Observing this particular OBSSR working group's recommendation in a planned mixed methods research study can contribute significantly toward increasing that study's rigor and quality. Accordingly, research investigators should provide explicit details regarding their proposed mixed methods research study design, including a clear rationale for their design choices, as linked to the purpose of their proposed study.

Mixed methods research designs. Hanson and colleagues have identified and described nine major *mixed methods designs* (Hanson, Creswell, Plano Clark, Petska, & Creswell, 2005), using nomenclature originally introduced by Morse (1991). Overall, the two distinct types of mixed methods research designs are (a) *concurrent designs*—designs in which qualitative and quantitative data are gathered within a single stage and (b) *sequential designs*—stagewise designs in which one form of data is gathered in stage 1, followed by stage 2, within which data from the other form are collected.

Concurrent designs. The three *concurrent* mixed methods designs identified by Creswell, Plano Clark, Gutmann, and Hanson (2003) are (a) *concurrent triangulation*, (b) *concurrent nested*, and (c) *concurrent transformative*. In each design, quantitative and qualitative data are collected during the same stage, although priority may be given to one form of data over the other. A major purpose of *concurrent triangulation* designs is to use both qualitative and quantitative data to generate a greater depth of understanding of existing relationships among variables of interest. As noted, in *concurrent nested* designs, both qualitative and quantitative data are collected during the same stage, although one form of data is afforded greater importance over the other (Creswell et al., 2003). *Concurrent transformative* designs are theoretically driven and designed to promote social change or advocacy while also used to generate support for a social perspective of interest (Hesse-Biber, 2012).

Sequential designs. In *sequential designs*, either the qualitative or quantitative data form is collected in an initial stage, followed by the collection of the other data form during a second stage. Within this typology, the three types of *sequential* mixed methods designs are (a) *sequential exploratory*, (b) *sequential explanatory*, and (c) *sequential transformative*. Each type has two subtypes—"*equal* priority"

and "*unequal* priority"—regarding the relative importance of their qualitative and quantitative components.

Conceptual Challenges in Conducting Scientific Mixed Methods Research

Fundamental Design Considerations and Imperatives Dissemination Models

This section presents four major mixed methods research design issues and challenges to consider in the design of rigorous mixed methods research studies: (a) effects of measurement and decontextualization, (b) QUAL and QUAN integration, (c) triangulation, and (d) sample size considerations. In an overview of the decision-making process in the design and conduct of a rigorous mixed methods study, Onwuegbuzie and Teddlie (2003) originally identified 12 important research design "decisions." Among these, four decisions of major importance are (a) deciding *the study purpose*, whether to conduct *between-methods triangulation* to cross-analyze a phenomenon (using a *concurrent research design*) or to conduct *initiation* (using a *sequential design*) to use stage 1 evidence to inform stage 2 of the study; (b) adopting a *variable-oriented* versus a *case-oriented* approach, whereby the variable-oriented approach is typically quantitative in form and the case-oriented approach is typically qualitative in form; (c) conducting an *exploratory* versus a *confirmatory* approach, that is, obtaining an initial view of a phenomenon (*exploratory*) or testing a hypothesis (*confirmatory*); and (d) deciding on the type of QUAL and QUAN data relationship, that is, whether the QUAL and the QUAN components will be equal in importance or whether the importance of one data form will be emphasized over the other (e.g., [qual → QUAN] or [QUAL → quan]) (Onwuegbuzie & Teddlie, 2003).

Measurement and Decontextualization

Measurement and coding are fundamental scientific operations that introduce a one-to-one correspondence between a real-world attribute and a form of data coding, whether quantitative or qualitative. The conversion of an "ecological event" into coded form generally "detaches" or *decontextualizes* that event from its natural environmental contexts. Also, such encoding is typically *reductionistic*, in the sense that it simplifies information as drawn from its more complex environmental contexts (Castro et al., 2010). Recently, the conversion of textual data into numeric form (i.e., "quantifying data") has gained greater acceptance among qualitative analysts, given that quantification offers certain advantages that include (a) ease in

identifying patterns, (b) affording analytic integrity, and (c) allowing hypothesis testing (Onwuegbuzie & Teddlie, 2003, p. 355).

Importance of Genuine QUAL and QUAN Integration

In principle, producing full integration between qualitative and quantitative data forms constitutes a major aim of mixed methods research, although this persists as a major research design challenge (Bryman, 2007). A seamless QUAL and QUAN integration is attainable from an a priori integrative plan that is implemented across all study components (i.e., conceptualization, data gathering, data analysis, and interpretation). Thus, the quality of a mixed methods research and its QUAL and QUAN integration begins with an integrative conceptualization of the study design and its data analytic plans (Castro & Coe, 2007). Regarding this issue, Bryman (2007) has stated that:

> the matter of how to present mixed methods findings in such a way that the quantitative and the qualitative findings are genuinely integrated, rather than standing as separate spheres or barely referring to each other, has not been touched upon to any significant extent in the burgeoning literature in the field. In genuinely integrated studies, the quantitative and the qualitative findings will be mutually informative (pp. 20–21).

Issues Involving Triangulation

A frequently invoked concept in qualitative analysis and in some mixed methods studies is the procedure of "triangulation." Triangulation involves the use of two or more methods to assess agreement (concordance) or disagreement among elements of an observation (Mertens, 2012). Triangulation in mixed methods research can be examined from broad to narrow conceptualizations. The narrow conceptualization, which some scholars brand as a "positivist" approach, involves using "triangulation as a validity tool." Under this narrow approach, when results from a second data form (often qualitative) converge with results from a first data form (often quantitative), then validation is said to occur (i.e., establishing convergent validity) (Fielding, 2012). A broad conception of triangulation is described by a feminist or interpretive approach whereby triangulation is used as a "dialectical process," which involves comparing and contrasting patterns of data for their convergences and divergences. This approach seeks to attain a more nuanced and deeper understanding of research findings by placing "disparate findings in dialogue with one another" (Hesse-Biber, 2012, p. 139). Under this broader conception, divergent or disparate results are just as important as convergent results given that emerging contrasts reveal new information that offers a more complex understanding of events under study, as this dialectic process also uncovers new and hidden knowledge (subjugated knowledge) (Hesse-Biber, 2012, p. 139).

Many studies typically conduct what may be described as, "informal triangulation." Such triangulation is invoked when qualitative and quantitative data sources

are regarded as concordant when they appear to be in agreement. Certain qualitative procedures, such as a "case-ordered effects matrix" or a "case-ordered predictor-outcome matrix," have been developed as aids for assessing such concordances (Miles & Huberman, 1994). These methods can provide an "initial glimpse" into real-world associations, although they remain *exploratory* in form rather than *confirmatory*.

From the narrow perspective, a more precise approach to triangulation involves the use of reliable measurement and planned rigorous methodologies as attainable by using conventional correlation analyses (Cohen et al., 2003). For example, one can examine the strength of association between measured quantitative data and converted qualitative data (i.e., qualitative textual data that have been reliably converted into numeric form). This approach can also reduce the cultural bias of extant measures as well as investigator bias (Mertens, 2003) and has been used in cross-cultural psychology (Triandis, McCusker, & Hui, 1990). As one example, Castro and Coe (2007) examined the strength of association between measured numeric variables and numerically converted qualitative text narratives (Castro et al., 2010) that were obtained from in-depth interviews. For example, significant associations were observed between the quantitative Acculturation Scale and the qualitative-converted "thematic variable" of Family Unity ($r = -0.32$, $p < 0.01$) and between the quantitative Family Traditionalism scale and the qualitative-converted thematic variable of Small Town Life is Better ($r = +0.27$, $p < 0.05$). Also, among the set of qualitative-converted thematic variables, a significant negative correlation was observed between the thematic variable of "Mucho Privilege" and "Family Unity" ($r = -0.26$, $p < 0.05$), and a significant positive correlation was observed between the thematic variables of "Values Traditions" and "Resists Change and [is] Old-Fashioned" ($r = +0.47$, $p < 001$).

Issues in Obtaining an Optimal Sample Size

A major research design challenge in mixed methods research involves capturing and integrating true aspects of the qualitative and quantitative traditions while reconciling major conflicts that can emerge from these research traditions (Denzin & Lincoln, 1994; Guba & Lincoln, 1994). As one example, these differing traditions often present differing perspectives regarding the number of research participants who constitute an optimally sized sample. This sample size quandary may be one of the areas in which the purely qualitative and purely quantitative traditions clash most when pondering considerations for attaining an optimal sample size.

Qualitative approaches to sample selection and size. From the qualitative tradition, needed sample size involving the number of cases to be examined can be determined by the criterion of *saturation*. The saturation criterion argues that case sampling can end when information attained becomes "redundant," that is, when producing repetition of information that is observed across new cases. Once the investigator "hears the same story" from a sequence of cases, saturation is said to occur, and it is thus "senseless" to continue recruiting additional cases. From a

purely qualitative perspective, the notion of saturation makes intuitive sense. By contrast, from the quantitative and mixed methods research perspectives, and particularly from the integrative mixed methods perspective (Castro et al., 2010), the saturation criterion introduces significant research design problems. The saturation criterion typically biases sample selection size toward the smallest viable number of cases. It thus introduces data analytic limitations when attempting to conduct quantitative data analyses with a sample having very few cases.

Another problem with the saturation criterion is "What constitutes the same story?" This point is seldom aptly described. Often, limited detail is given on aspects of a story that exhibit redundancy and thus produce "saturation." Although some aspects of a participant's story may be repeated, thus suggesting the occurrence of "saturation"; other aspects may not. In terms of the number of cases, for a qualitative study, an optimal number of cases typically ranges from 12 to 40, although in one study, the saturation criterion was used to sample only four cases (Montaghan, Sanders, Kelly, Cogen, & Streisand, 2011).

From the qualitative perspective, and as a practical consideration, analyzing text narratives is labor intensive. Accordingly, every new case extends the time and effort involved in data preparation, analysis, and interpretation. It thus makes sense to gather the fewest number of necessary cases. As such, qualitative studies are driven toward obtaining small sample sizes. Unfortunately, in rigorous mixed methods research, beyond issues of statistical power, a small number of cases can limit the design's internal validity; a very small sample introduces a selection bias that confounds the ability to attribute an observed effect to the factor under study. A small group size can also compromise external validity by limiting the representativeness of the participants sampled to the larger group of interest, thereby also limiting the generalizability of the results. And as noted, in practice, the saturation criterion is often applied in a subjective manner, based largely on global judgments regarding a perceived point of redundancy.

Quantitative considerations in determining adequate sample size. By contrast, sample size determination in quantitative research is driven by concerns over the power to detect an effect, based on an a priori determination of a "minimally interesting effect size," also taking into account the levels of allowable Type I and Type II errors (Cohen et al., 2003; MacCallum, Browne, & Sugawara, 1996; MacCallum, Widaman, Zhang, & Hong, 1999). This quantitative perspective argues that "too few cases" indicate that a study is underpowered, thus compromising the capacity to detect the effect of interest and to conduct robust statistical tests of significance. Quantitative studies are thus driven toward gathering larger sample sizes, typically ranging from 100 to 1,000 cases (Gelo, Braakman, & Benetka, 2008; Yoshikawa et al., 2008).

In parallel with issues of subjectivity as introduced in the qualitative assessment of saturation, the quantitative analysis of power and sample size may appear rigorous, although power calculations are often based on an investigator's use of plausible but at times not entirely verifiable assumptions (e.g., the expected effect size). Thus, judgments of an expected *effect size* are often overly optimistic when research investigators attempt to justify a desired sample size. Subsequently, the

investigator may conduct numeric and formula-driven power calculations that can generate accurate estimates of power, provided that the stated assumptions are correct. If not, then the power calculations are flawed.

Conclusions

In conclusion, the design of a rigorous mixed methods study must often reconcile certain competing or conflicting QUAL and QUAN imperatives, which include identifying an optimal sample size for a proposed mixed methods research study. Identifying an optimal mixed methods design, as illustrated here by the identification of an optimal sample size, is best done by careful consideration of several research design alternatives that include (a) the purpose of the study (Onwuegbuzie & Teddlie, 2003), (b) the research question and/or hypotheses to be tested, and (c) the related quantitative and qualitative data analyses to be conducted. This also involves considerations regarding (d) the depth of analysis of the text narratives as required per case and throughout the study.

Review of Contemporary Mixed Methods Studies with Applications for Prevention Science

This section examines the strengths and weaknesses of 14 recent research studies to highlight major issues in the design of mixed methods research studies, as applicable to prevention science research. Table 20.1 presents these 14 studies, which used qualitative and quantitative data forms to inform the planning, design, and/or evaluation of a prevention intervention. These studies were identified from an online literature search for the period from 2007 to 2012 using the keywords "mixed methods," "prevention," and "intervention" within the "abstract" search field. We accessed these studies using the PsycINFO database. From 57 studies identified in this search, we included 13 studies: (a) those that met these search criteria, (b) those that were also available as full text documents, and (c) those that also consisted of an empirical intervention study. In other words, we excluded articles that consisted solely of a review of the literature or a discussion of research issues. We also added one recent study that appeared in the journal *Prevention Science*, a study that used a mixed methods approach to evaluate a prevention intervention (O'Brien et al., 2012). Table 20 thus includes a total of 14 studies, 4 *concurrent,* and 10 *sequential* mixed methods studies.

Table 20.1 Mixed methods studies applied to prevention research

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
	Concurrent designs				
1	Team: Chakrapani et al. (2011)	Aim: To identify sociocultural contextual factors among injection drug users in Northern India	MM-design: A [QUAN + qual] design	Application: Identified key enabling factors and modifiable contextual factors (e.g., police harassment) that can promote safe needle use among IDUs	Strength: Use of multiple qualitative methods to elicit contextual information from participating IDUs
	Topic: Social–structural contexts of needle and syringe sharing behaviors among HIV-positive drug users	Sample:75 HIV-positive injection drug users (IDUs), both males and females	Described by the authors as an "integrated mixed methods investigation" (p. 2)		Partial weakness: No formal integration of qualitative and quantitative content; some concordance reported in responses from IDUs and key information
	Location: India	Sample: In-depth interviews with 4 male and 3 female IDUs and 2 key informant physician interviews	Structured survey with $N = 75$ HIV-positive IDUs		Weakness: Lacks a rationale for the research design and methods used and lacks a strategic and integrative design to allow a unified "multiple methods" and integrative analysis of study results
		Sample: Two focus groups with 9 men, and the second with 8 women	Conducted two focus groups ($n = 17$), seven in-depth interviews, and two key informant interviews; no rationale for this sampling of participants		
		Approach: A mixed methods study consisting of a cross-sectional quantitative survey, focus groups, and in-depth interviews			
2	Team: Lewis et al. (2008)	Aim: To examine social and cultural aspects of gender relations and sexual	MM-design: A concurrent [QUAL + QUAN] design	Application: Qualitative description of five forms of violence (physical,	Strength: Community participants inform the

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
		behavior as factors in HIV risks among women in Papua New Guinea		emotional, financial, sexual abuse, or social isolation) provided additional details to inform the development of a violence prevention intervention	development of the interview protocol
	Topic: Factors affecting domestic violence toward women and its association with HIV risks	Sample: $N = 415$ women	A "mixed methods approach" (p. 183)	Qualitative responses elicited women's preferences regarding needed components for designing a prevention intervention (i.e., support, educational, and sociocultural services)	Partial weakness: No formal qual–quan integration; incorporated a partial integration of qual–quan data consisting of open-ended responses to elicit details in response to the quantitative items that asked about violence, when these were answered affirmatively
	Location: New Guinea	Approach: One-time face-to-face interviews using structured quantified items and open-ended qualitative questions	Quantitative analyses examined subgroup differences between HIV-positive and HIV-negative participants Qualitative analysis encoded emergent themes		Weakness: Little detail provided regarding thematic analysis of textual data
3	Team: Tummala-Narra et al. (2011)	Aim: To understand the effects of racism	MM-design: A concurrent [QUAL + QUAN] design	Application: Mixed methods model analyses can identify moderator	Strength: Examined race and ethnicity using quantitative and qualitative data

		experienced by Asian Indian adults		and/or mediator effects that influence specific mental health outcomes and can identify components of change for informing the development of a prevention intervention	to understand the effects of cultural variables on positive mental health outcomes Weakness: Internet sampling yielded a select and biased sample Weakness: Lacked a formal integration of qualitative and quantitative data
	Topic: Asian Indian youth responses to discrimination Location: United States	Sample: $N = 102$ participants sampled using an online survey Approach: Examined first- and second-generation Asian Indians; examined how racial and ethnic identity may mediate the effects of racial stress on self-esteem and coping	The authors describe using a "mixed methods approach" (p. 205) Used a thought-listing technique to elicit *situations* and *reasons* in which race and ethnicity were salient Thematic analyses were used to create categories of *situations* and *reasons* that promote the salience of race and ethnicity Quantitative analyses tested models of racial stress as mediated by racial identity as influences on self-esteem and coping Quantitative analyses examined generation status and racial socialization as moderators of racism on self-esteem and on collectivistic coping		
4	Team: Hanbury et al. (2011)	Aim: Based on the theory of planned behavior (TPB), examined interventionist behaviors in the delivery	MM-design: A concurrent [QUAL + QUAN] design	Application: Examined factors that influence staff adherence in the delivery	Strength: Raised questions regarding the appropriateness for a broader systemic analysis of

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
		of a suicide prevention intervention		of a prevention intervention	using the TPB, which is an individual decision-making theory
	Topic: Mixed methods approach to conduct a temporal analysis of behavior change outcomes and the process of change; examines both efficacy and reasons for efficacious outcomes	Sample: Originally, $N = 93$ health-care professionals although only $n = 50$ (54 %) submitted the survey	The authors assert using a "mixed and multiple methods of evaluation"(p. 291)	In a qualitative process evaluation, narrative texts, in the form of anonymous quotes, identified staff beliefs about adherence and can inform approaches for tailoring the intervention	Weakness: Low response rate to the staff survey (i.e., 54 %)
	Location: England	Approach: Program evaluation using quantitative and qualitative data to identify barriers to staff adherence when delivering an intervention protocol; aided in identifying why the intervention was or was not successful	Qualitative survey and key informant interviews conducted with 3 community psychiatric nurses and 2 primary care facilitators who monitored adherence to the intervention protocol	Can identify practical barriers that limit the feasibility of a prevention intervention; aid in developing practical implementation strategies	Weakness: Mentioned conducting a "content analysis" of textual narratives; however, details were not provided on how it was conducted
			Time series analyses, mediational analysis, and qualitative process analysis were used to assess effectiveness of intervention to promote staff adherence		Weakness: A very small sample, where only $N = 5$ interviewees were used to obtain qualitative process data to assess intervention delivery Weakness: No formal integration of qualitative and quantitative data

	Sequential designs				
5	Team: Armstrong and Boothroyd (2007)	Aim: To identify predictors of emotional well-being	MM-design: A [QUAN → QUAL] design(repeated yearly)	Application: This approach offers a deeper understanding of risk factor effects on emotional well-being, informing the development of strategies to attenuate risks	Strength: Findings informed the development of a tertiary mental health intervention
	Topic: Factors to inform prevention strategies for promoting emotional well-being among at-risk adolescent girls	Sample:125 at-risk adolescent girls (ages 13 to 17) whose families were receiving welfare	A "4-year, mixed methods study" (p. 435)	Qualitative interviews following the quantitative logistic regression analyses generated illustrative examples of effects identified in the quantitative analyses; this approach can inform the design of prevention interventions	Partial weakness: Lack of formal integration, although an informal link was created between results from the quantitative analyses and those of the qualitative analyses
	Location: United States	Approach: Quan findings were informed by subsequent qual results	A 4-year longitudinal study with yearly qual–quan assessments A one-group longitudinal repeated measures design Stage 1 (repeated yearly), a survey of $N = 125$ mothers and daughters Stage 2 (repeated yearly) in-depth interviews with 20 girls drawn from the original sample; open-ended interview "explored in greater detail many of the same factors" assessed in stage 1	Mixed methods findings were used to "propose prevention activities using a tertiary mental health prevention intervention framework" (p. 435)	Weakness: Study design consisted of a one-group repeated measures study without a control group

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
6	Team: Bowleg et al. (2012)	Aim: To conduct an in-depth "intersectionality analysis" of interrelationships among race, gender, and SES, as influences on participants' experiences of social inequality	MM-design: A [QUAL → QUAN] design	Application: Stage 1 of this study provides an exploratory analysis of complex interactions involving race, gender, and SES, as basis for informing and developing content and a research approach for a stage 2 study and the development of a future HIV prevention intervention	Strength: Conducted an in-depth analysis of the experiences involved in being a Black man
	Topic: *Intersectionality*—an analysis of the intersecting and synergistic effects of low socioeconomic status (SES), Black race, and male gender on social inequality, thus analyzing "interlocking systems of oppression among Black men" (p. 8)	Sample: $N = 30$ MSM ages 18 to 44	"The initial qualitative stage of a larger mixed methods HIV prevention study" (p. 1)		Strength: Used branching logic to identify logical positive and negative aspects of emerging themes
	Location: United States	Approach: Stage 1 was a qualitative study	Stage 1 qualitative study identifies four major themes		Strength: One atypical case "gave voice" to challenge the assumption of putative adverse influences of racial identity and discrimination; this case introduces alternative conceptualizations regarding the experiences of MSMs

			Stage 2, a larger planned study, to be conducted in the future		Weakness: The manner in which questions were framed may have biased the types of responses elicited Weakness: Representativeness of themes and concerns expressed by this sample of Black MSMs has limited generalizability of results as applicable to other Black MSMs from this same demographic niche
7	Team: Capella et al. (2011)	Aim: To use a participatory social action research approach to inform the local adaptation and implementation of a BRIDGE (Bridge Education and Mental Health Urban Schools) program	MM-design: A [QUAL → QUAN] design	Application: Systematic gathering of field data provided a general overview of major implementation issues and challenges, as this informed the development of a school-based intervention	Strength: Collective input from school and community key informants identified program weaknesses and identified issues and concerns to aid in "local adaptation," program refinement, and a more effective implementation, as these improvements can increase the intervention's relevance and sustainability
	Topic: Use of participatory social action research to	Sample: Stage 1 (Phase 1) had a diverse sample of	"Phase I involved analyses of mixed method	Two focus groups aided in refining the program	Strength: The use of consensual qualitative "open

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
	inform intervention development	($N = 36$) participants including administrators, teachers, supervisors, and agency clinicians	implementation data from a pilot experimental trial of the adapted program" (p. 486)	manual and in identifying barriers to implementation	coding" identified *emergent themes*; a team review procedure was used to resolve discrepancies
	Location: United States	Sample: Stage 2 (Phase 2) had ($N = 58$) diverse participants including teachers, mental health professionals, and consultants	Stage 1: The qualitative data consisted of field notes and focus group qualitative data, "to inform refinement to training, content, delivery and supervision" (p. 406)		Strength: Stage 1 identified concerns expressed by key informants, thus eliciting "consumer preferences" that inform modifications in stage 2 in response to consumer preferences
		Approach: A complex two-stage formative evaluation was conducted to build collaboration between mental health professionals and teachers and to refine a BRIDGE program for effective collaboration among staff from local schools	Stage 2: A pilot experimental trial was conducted of the revised and adapted BRIDGE program, as this involved the "analysis of mixed methods implementation data from the pilot experimental trial of the adapted program" (p. 486)		Weakness: Stage 2 did not consist of a formal controlled trial Weakness: An informal and non-systematic "impressionistic integration" was conducted across stages
8	Team: Cianelli et al. (2008)	Aim: To understand how sociocultural factors, including *machismo* and *marianismo*, contribute to sexual violence and	MM-design: A [QUAL → QUAN] design	Application: In-depth qualitative analyses can generate culturally rich information that can enhance the cultural	Strength: Presented rich detail regarding low-income women's perceptions of gender role expectations and

	HIV risks and are barriers needing elimination in the development of an HIV prevention program		relevance and inform the development of an HIV prevention intervention	norms, as risk conditions for HIV infection as well as operating as barriers to women's empowerment
Topic: HIV prevention among low-income Chilean women	Sample: In stage 1, sampled 50 low-income Chilean women	Described by the authors as a "mixed methods study" (p. 297)		Weakness: Report low concordance between outcomes derived from the quantitative survey and from the qualitative content analyses
Location: Chile	Sample: In stage 2, 20 of these women were selected for in-depth interviews	A one-group pre–post program evaluation using quantitative measures in stage 1, followed in stage 2 by text narratives about the program obtained from the focus groups		Weakness: Few details were given regarding the analyses of text narratives; it is only stated that a "content analysis" was conducted to "recognize, code and categorize patterns from text data" (p. 300)
	Approach: A one-group pre–post program evaluation was conducted using quantitative measures in stage 1, followed in stage 2 by qualitative text narratives generated from focus groups			Weakness: No formal integration of qualitative and quantitative data was conducted
	Approach: In stage 1, a survey was conducted to assess HIV-related risk perceptions. In stage 2, 20 in-depth interviews			

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
		explored gender role norms (*machismo* and *marianismo*) as factors in sexual violence and HIV risks			
9	Team: Coombes et al. (2009)	Aim: To conduct an exploratory study of change in adolescent knowledge and refusal skills against the use of drugs and alcohol, regarding the Strengthening Families program	MM-design: A sequential [QUAN → QUAL] design	Application: Post-intervention reports from intervention participants were used to provide additional contextual information regarding participants' views of the intervention, to identify potential sources of efficacy, and to aid in improving the original intervention	Weakness: The investigators identified several research design weaknesses that included a small, biased and non-probability sample; the absence of a control group; the lack of valid and reliable measures; and the absence of a long-term follow-up observation
	Topic: Evaluation of a Strengthening Families program based on its implementation in an urban community in Northern England	Sample: A total of 58 families from an original sample of 70, who previously completed the Strengthening Families program	The authors describe using a "mixed methods design blending both quantitative and qualitative data" (p. 41)		Weakness: All thematic content analyses of focus group data were conducted without reporting details of that methodology
	Location: England	Sample: Included a post-intervention sample of 15 facilitators (program delivery staff) in focus groups to comment on	The study was described as consisting of a "mixed methods design" that "blends" the use of quantitative and qualitative data		Weakness: Separate analyses of quantitative and qualitative data were followed by a reported "synthesis" of the main findings from both

		analyses of intervention results			approaches; however, no formal methodology was reported regarding procedures for "blending" the qual–quan data
			Stage 1: A one-group pre–post design was conducted with data gathered using a quantitative survey Stage 2: Two focus groups (5 families each) were sampled from 10 of the 58 families who had participated in the prior survey Stage 2: Facilitator focus group; 3 focus groups (with 5 participants per group) were conducted, involving 15 of the 30 facilitators who delivered the program		Weakness: Pre–post data provided participants' reports of improvements on certain aspects of family function. However, many of the conclusions regarding intervention efficacy relied on qualitative self-reports obtained from care givers (parents) and children
10	Team: Montaghan et al. (2011)	Aim: To examine aspects of parental support among parents of children who have type 1 diabetes; assessment of parental satisfaction and their reactions to intervention components	MM-design: An exploratory [QUAL → QUAN] design	Application: Quan assessment of participant satisfaction with the parent education intervention	Weakness: *Saturation* criterion used that resulted in a sample of only 4 cases that were used to identify themes about the intervention

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
	Topic: Evaluation of the Supporting Parents program	Sample: $N = 12$ parents of a child having type 1 diabetes	"Presents mixed-methods program evaluation data" (p. 868)	Select cases ascertain aspects of intervention effects and parental satisfaction with the intervention	Weakness: No formal integration of qual–quan data
	Location: United States	Approach: A mixed methods program evaluation of a five-session intervention to promote parental mastery of tasks for helping their diabetic child	A satisfaction survey with $N = 12$ parents An in-depth assessment with $n = 4$ of program completers	Identifies critical components to aid in intervention refinement	
11	Team: Nelson and Tom (2011)	Aim: To evaluate the effects of a Gathering of Native Americans (GONA) intervention for reducing youth risks for HIV/AIDS, hepatitis, and substance use	MM-design: A sequential explanatory [QUAL → QUAN] design	Application: Qualitative analyses provided an "insider's perspective on phenomena not captured by questionnaire and to employ community participation in analyzing qualitative results and assessing the strengths and weaknesses of the intervention" (p. 351)	Strength: A sensitivity to cultural issues affecting Native Americans
	Topic: Evaluation of a substance abuse and HIV prevention intervention for urban Native American youth	Sample: $N = 100$ youth ages 13–18	"A mixed methods outcome evaluation" (p. 349)		Strength: Developed open-ended questions to elicit participants' post hoc explanations about results of tests that examined pre–post changes from stage 1
	Location: United States	Approach: A one-group pre–post design was used to assess the efficacy of a prevention intervention for changing HIV and	A one-group pre–post design to examine changes in specified outcome measures		Weakness: A one-group pre–post design was implemented, although it was compromised by sampling bias and

		drug use-related changes. These changes targeted knowledge, risk perceptions, sexual self-efficacy, sexual risk behaviors, and ethnic identification	Conducted a thematic analysis of qualitative data		considerable attrition at the posttest Weakness: Little detail was provided regarding methods for conducting the qualitative thematic analyses Weakness: The use of simple statistical tests regarding pre–post changes, where these analyses were also limited by the absence of a control or comparison group
12	Team: O'Brien et al. (2012)	Aim: To identify and understand in-depth the factors that promote attrition and retention in the NFP program	MM-design: A sequential [QUAN → qual] design	Application: The multivariate model identification of high and low attrition sites served as the basis for conducting an in-depth thematic analyses of factors related to attrition and allowed an in-depth understanding of modifiable factors that aid in enhancing the intervention and its implementation by reducing participant attrition	Strength: The core study provided the basis for the design and implementation of a qualitative sub-study to further understand factors that contribute to and that discourage participant attrition
	Topic: A study of the factors (quantitatively identified) and their explanations (qualitatively identified) that were associated with levels of participating mothers' departure from the Nurse–Family Partnership (NFP) program	Sample: In stage 1, a total of 10,367 participating mothers recruited from 66 NFP implementation sites	The qualitative analysis from these high and low attrition sites helped to "gain insight into how nurses in these two groups of sites approached the retention of participants" (p. 220)		Partial weakness: Quantitative analyses were informed of the selection of sites having *highest* and *lowest* rates of participant attrition; a thematic analysis of these factors yielded explanations about the quantitatively identified factors; a

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
					partial level of integration of quan–qual data was attained
	Location: United States	Approach: In stage 2, from the highest ($n = 5$) and lowest ($n = 6$) attrition sites, conducted a qualitative thematic analyses of responses from staff nurses and supervisors regarding their views of how to retain the participating moms	The core NFP program aims to (a) improve pregnancy outcomes, (b) improve the infant's health and development, and (c) increase mother's personal development		
13	Team: Simoes et al. (2008)	Aim: To identify risk and protective factors among delinquent youth	MM-design: A [QUAN → QUAL] design involving "construct triangulation"	Application: Stage 2 focus group outcomes could be used to identify concordance with results from the SEM analyses	Strength: Concordance was reported in the identification of major risk and protective factors identified by the stage 1 quantitative data and by the stage 2 qualitative data as elicited from participating youth
	Topic: Analysis of risk and protective factors for juvenile delinquency by using quantitative and qualitative methods	Sample: A total of 275 juvenile offenders ages 11–18 in stage 1 and a purposive sample of 24 of these youths in stage 2	"The present study was developed using mixed methods" (p. 389)		Strength: The SEM model analysis identified major risk and protective factors (e.g., delinquent friends, poor parent–child relationships, good family communications),

					which were then used as topics in focus groups to elicit youth views regarding the dynamics of these factors
	Location: Spain	Approach: A 2-stage study to build a structural equations model of delinquency (Study 1), followed by focus group feedback on model constructs (Study 2) to add depth of explanation to model constructs	In stage 1, conducted a quantitative structural equations model (SEM) analysis of 275 juvenile offenders to develop an explanatory model of youth delinquency		Weakness: Little detail was provided regarding the use of the focus group
			In stage 2, conducted four focus groups with $n = 24$ youth offenders to understand their perceptions regarding risk and protective factors for delinquency		Weakness: A formal integration of results derived from qualitative and quantitative data was not conducted
14	Team: Wilkerson et al. (2011)	Aim: To evaluate intervention components that would be efficacious in decreasing HIV risks	MM-design: A [QUAN → QUAL] design	Application: Expanding a self-report qualitative response with an open-ended probe following that item, thus adding depth of response and a context to that qualitative response—an "item response contextualization"	Strength: A theory-driven study, with sound randomization and three observation points (pretest, posttest, and 3-month follow-up)

(continued)

Table 20.1 (continued)

No.	Team of investigators, topic, and location	Aim, sample, and approach	Mixed methods design (MM-design)	Applications to prevention science	Design strengths and weaknesses
	Topic: Analysis of *critical self-reflection* as a "discovered factor" that can facilitate changing HIV risk perceptions among men who use the Internet to have sex with men	Sample: $N = 560$ participants randomized to an online risk-reduction intervention or to a wait-list control condition	A "mixed-methods approach" (p. 15)	This approach is useful in informing the design of an HIV prevention intervention; results suggested the utility of *critical self-reflection* for changing sexual beliefs and perceptions involving HIV risks	Strength: An "item response contextualization" was used to contextualize quantitative responses
	Location: United States	Approach: A randomized controlled trial (RCT) to develop and test an Internet-based HIV prevention intervention for US men who use the Internet to seek sex with men. Qualitative thematic analyses explore cognitive mechanisms underlying the behavior change observed from the prevention intervention	A controlled randomized (two-group) trial		Strength: Qualitative responses were collected systematically and coded thematically to identify major themes
			Baseline survey, posttest, and 3-month follow-up		Strength: Comments about reflections yielded an emergent theme that led to the identification of *critical self-reflection assumptions* (CSRAs) as an important construct
			At posttest, open-ended questions asked about participants' likes, dislikes, and suggestions for improving the intervention		Strength: Qualitative quotes were used to illustrate reasons for change, as linked to the quantitative responses
			At 3-month follow-up observation, open-ended questions asked about intervention's effects on		Partial weakness: Formal integration was not conducted, although some quan–qual

thoughts, attitudes, and behaviors; control group asked this about the survey items		integration was attained for informing change effects; used an "item response contextualization" by eliciting qualitative text narratives on avowed reasons for behavior change

Strengths and Weaknesses of the Identified Studies

These studies varied in their methodological strengths, and all were included to present a diverse sampling of recent studies that professed to use a "mixed methods" approach, broadly defined, and to use it to develop, implement, and/or evaluate a "prevention intervention," also broadly defined.

Overview of studies purporting to conduct mixed methods research. An overview of these studies reveals that 7 of these 14 studies were conducted outside the United States, suggesting the international appeal of mixed methods research approaches. Some studies were described as a "mixed methods" study based solely on a gathering of numeric (quantitative) data and textual (qualitative) information, despite exhibiting little or no integration of results (Bryman, 2007), as obtained from these two data forms. In fact, it appears that none of these 14 studies was designed a priori to allow a full integration of the qualitative and quantitative components.

Major Limitations in Contemporary Mixed Methods Studies of Intervention Outcomes

The following sections present five major design limitations that compromised, to varying degrees, the internal and/or external validity of these putative mixed methods studies. These limitations are (a) an unspecified mixed methods design, (b) weak or unspecified purposive sampling, (c) the use of a one-group experimental design, (d) the use of impressionistic integration, and (e) a methodology that permitted only partial qualitative–quantitative integration.

An Unspecified Mixed Methods Design

Several investigators referred to their study as a "mixed methods approach," although as noted, none of these 14 studies explicitly specified the type of mixed methods research design that was employed. Three of these studies generally referred to their use of a "mixed methods approach" within a more rudimentary program evaluation study (Lewis, Maruia, & Walker, 2008; Montaghan et al., 2011; Tummala-Narra, Inman, & Ettigi, 2011; Wilkerson, Danilenko, Smolenski, Myer, & Rosser, 2011). Other studies alluded generally to their use of a "mixed methods" methodology (see column 3 in Table 20.1). The study design features that would qualify a given study as a mixed methods study varied considerably. A significant question in this regard is, what research design elements constitute the essential features of a genuine mixed methods study? Does solely collecting qualitative and quantitative data constitute a sufficient criterion for classifying a study as being a "mixed methods study"? Conversely, how essential is creating a genuine qualitative–quantitative

integration to classify a study as being a true "mixed methods" study? Among these studies, given the absence of an explicit statement regarding the particular type of mixed methods research design used, we added a "Mixed Methods Design (MM-Design)" designation and its nomenclature, which is presented in column 3 of Table 20.1. This constitutes our best indication of the mixed methods design likely used, based on the authors' description of their research approach.

Weak or Unspecified "Purposive Sampling"

A recurring weakness in several sequential QUAN → QUAL research studies involves sampling limitations in the qualitative component involving small samples ($n < 30$), in the total sample of focus group participants, and/or in the total sample of key informants, along with the absence of a clear rationale for sample selection (Armstrong & Boothroyd, 2007; Chakrapani, Newman, Shunmugam, & Dubrow, 2011; Cianelli, Ferrer, & McElmurry 2008; Coombes, Allen, Marsh, & Foxcroft, 2009; Hanbury, Wallace, & Clark, 2011; Montaghan et al., 2011) (see Table 20.1). Some investigators reported using a "purposive sampling approach" yet provided no details regarding the criteria used for this type of sampling. In addition, one study invoked the *saturation* procedure, thus orienting toward a smaller-sized sample (Montaghan et al., 2011). In both quantitatively focused and qualitatively focused studies, such small-sized samples can produce unstable, inaccurate, or misleading conclusions when informing the design of a planned prevention intervention.

A One-Group Experimental Design and Limited Internal Validity

Five of the ten sequential design studies (Armstrong & Boothroyd, 2007; Cianelli et al., 2008; Coombes et al., 2009; Montaghan et al., 2011; Nelson & Tom, 2011) conducted a simple program evaluation study that consisted of a one-group pre–post research design. These weaker studies clearly exhibited low internal validity and thus a low capacity to unambiguously assess the effects of the intervention. By contrast, two of these studies (O'Brien et al., 2012; Wilkerson et al., 2011) used a stronger design that had comparatively greater internal validity.

"Impressionistic Integration"

As has been noted, some investigators have alluded to conducting a "mixed methods" study, based primarily on their collection of qualitative and quantitative data, although doing so in "separate spheres," with these data forms "barely referring to each other" (Bryman, 2007). In this regard, these studies also lacked a formalized plan to integrate these data forms. These investigators also referred

explicitly to their efforts at "blending" or "synthesizing" their study results (Coombes et al., 2009). However, claims of such "blending" were accompanied with a lack of detail regarding methods of content analysis or the elicitation of thematic analyses, coupled with vague links to the quantitative data. These weak efforts at data integration may be described as "impressionistic integration" (Cianelli et al., 2008; Hanbury et al., 2011; Lewis et al., 2008; Nelson & Tom, 2011). Such impressionistic methods generated limited or sketchy evidence for use in informing the development of a planned prevention intervention. By contrast, a more carefully planned, rigorous, and "genuine integration" of QUAL and QUAN data (Bryman, 2007) could have generated more definitive and convincing results for informing the design of a new or planned prevention intervention. Among these studies, only the study by O'Brien et al. (2012) generated specific information for improving the design and implementation of the prevention intervention under study.

Partial Qualitative–Quantitative Integration and "Item Response Contextualization"

Four studies (Armstrong & Boothroyd, 2007; Chakrapani et al., 2011; Lewis et al., 2008; Wilkerson et al., 2011) conducted a data-gathering strategy that facilitated some level of integration of their qualitative and quantitative data. We are referring to this as "item response contextualization," in which participant responses to a quantitative item or question were followed immediately by an open-ended question to thus allow respondents to elaborate on their given response. We describe this approach as a "partial limitation" in these research designs. For a given response, this item format elicited an explanation and contextualization following the initial quantitative response, thus providing a direct QUAL and QUAN link that allowed a certain level of QUAN and QUAL integration. The construction of conventional structured interview items when followed with an open-ended question that allowed participants to elaborate on their initial response added contextual or explanatory details to the quantitative closed-ended items. If well designed, such composite QUAL and QUAN items can offer a partial level of QUAL and QUAN integration that can generate greater depth of information, which in turn can better inform the design of a new or improved prevention intervention.

Recommendations for Mixed Methods Research Studies in Prevention Science

Prevention science is a research field that is characterized by studies that exhibit high validity (both internal and external) and are designed to address various significant prevention science issues. Accordingly, for use in prevention science

research, a planned mixed methods study should adhere to the standards of evidence that guide the conduct of sound research in prevention science (Flay et al., 2005). Here it is recognized that in designing a rigorous mixed methods research study, the application of these prevention science standards of evidence requires an additional round of review involving considerations for addressing many of the aforementioned issues that involve balancing the research design imperatives emphasized by the qualitative and the quantitative research traditions. Regarding these prevention science standards of evidence, most of the mixed methods research studies that we examined revealed the need for greater detail in the description of the qualitative procedures and analyses used, as well as greater rigor overall in their research design planning, specificity, and organization. We have identified four significant weaknesses as observed across these 14 contemporary studies that purported to conduct "mixed methods" research. These four weaknesses were (a) an unspecified mixed methods design, (b) weak or unspecified purposive sampling, (c) the use of a one-group experimental design, and (d) the use of impressionistic integration. Regarding ways to address these identified weaknesses, we offer the following recommendations to guide the design of more rigorous mixed methods research studies, as these can contribute new and generalizable knowledge for the field of prevention science.

Recommendation 1: Specify the intended mixed methods design. In accord with the recommendation presented by the OBSSR working group report titled *Best Practices for Mixed Methods Research in the Health Sciences* (Creswell et al., 2011), it is recommended that a planned mixed methods study specifies and describes the type of mixed methods design to be implemented. This would aid in enhancing research design planning and could aid in increasing the rigor incorporated into a proposed mixed methods research study. Ideally, the design of a genuine mixed methods research study would use the mixed methods nomenclature that specifies the type of mixed methods design to be used, setting the stage for greater a priori planning and a more regimented QUAL and QUAN data collection, which could then facilitate the conduct of genuine mixed methods data analyses. Also, research investigators should provide a clear rationale for their choice of research design, as linked to the purpose of their proposed study.

Recommendation 2: Specify the sampling plan and methodological procedures for the qualitative component. Many mixed methods studies have essentially obtained a sample of convenience for their qualitative component, whereby often this consists of a very small sample (e.g., two focus groups or five key informants). Such very small samples typically constitute a research design weakness, in both qualitative and quantitative research designs. They often provide nonrepresentative and biased responses from which to inform the study's quantitative component and/or the design or refinement of a prevention intervention. Future investigators are encouraged to provide a clear rationale for their choice and number of participants to be selected into their sample and to increase the representativeness of their sample as linked to their larger study sample. This includes an attempt to draw a better defined and relatively larger sample rather than invoke the saturation criterion to obtain a smallest possible sample. In connection with greater specification in

the sampling plan, greater specificity is needed throughout the study and its components, such as by providing clear details on the procedures used in extracting thematic categories from text narratives (Castro et al., 2010).

Recommendation 3: Use a study design that has greater internal validity. Some investigators conducted a mixed methods study that consisted of a simple one-group pre–post research design. This simple program evaluation-oriented design generates some results, although introducing several known limitations in the study's ability to attribute observed outcomes to the intervention itself (i.e., in internal validity). This simple one-group pre-–post design is typically of limited utility for conducting sound research in prevention science. Accordingly, to the extent viable, future mixed methods research for prevention science should develop a core study design consisting of a controlled randomized trial or a well-operationalized multivariate model study, one that uses clear QUAL and QUAN procedures and an explicit integrative data analysis plan (Castro & Coe, 2007).

Recommendation 4: Maximize integration in accord with the study purpose. Among the 14 studies reviewed, a QUAL and QUAN integration was incorporated only in part or not at all, despite the fact that integration of QUAL and QUAN features constitutes a major aim in mixed methods research. Of course, the extent necessary of a full QUAL and QUAN integration will depend on the nature of the proposed study, that is, whether it serves as an *exploratory* study (typically requiring less integration) or a *confirmatory* study (typically requiring more integration). To realize the full potential of mixed methods QUAL and QUAN data gathering and analysis, even for exploratory studies, incorporating a requisite level of QUAL and QUAN integration will be important. Accordingly, we recommend the a priori design of mixed methods studies that purposefully "builds in" QUAL and QUAN integration across all major study components: conceptualization, data gathering, implementation, data analysis, and data interpretation. The ultimate aim is to maximize the yield of mixed methods research studies for advancing the field of prevention science.

Recommendation 5: Training and mentorship in mixed methods research. The design of a well-crafted mixed methods research study for prevention science research can be enhanced by mastery of both the quantitative methodological standards of evidence (Flay et al., 2005) and the principles of qualitative research. This involves skills development for resolving competing or conflicting qualitative–quantitative design issues to craft a well-balanced research design that can maximize gains afforded by qualitative and quantitative data forms. Accordingly, training to develop skills for designing and conducting "genuine mixed methods research" should feature a curriculum that builds an in-depth understanding and appreciation of the qualitative and quantitative research traditions (Hesse-Biber, 2010) and of their integration. Such training should include mentorship and exercises to build capabilities for working with the dynamic tensions that can arise in addressing emerging conflicts between these distinct research traditions (Denzin, 2012).

Furthermore, conducting mixed methods research is a team effort (Creswell et al., 2011). It engages a diverse group of research investigators who appreciate

working across disciplinary boundaries. Thus, among many important training goals and objectives, training in mixed methods research involves (a) leadership for exercising the role of the principal investigator who guides the mixed methods research team, (b) organizing and managing the infrastructure for conducting mixed methods research, (c) comfort and proficiency in the analysis of multiple types of data, and (d) an openness to collaborations with diverse scholarly disciplines (Creswell et al., 2011). The subject matter in prevention science presents many excellent opportunities for the application of mixed methods methodologies to more effectively address several important prevention science research issues to effectively address the many health inequities and other social problems, with the aim of generating more informative and efficacious prevention science research studies.

Summary and Conclusions

This chapter presented major issues and challenges in the design of rigorous mixed methods research studies that are applicable to prevention science research. It is recognized that not all mixed method research studies are *confirmatory* in purpose or study design, given that some studies are explicitly intended to be *exploratory* in nature. Nonetheless, we have argued that rigor in conceptualization, design, implementation, and data analysis is essential for the conduct of potent mixed methods research studies, as these can better contribute significant scientific research results that will better inform the design of prevention science interventions. In conclusion, mixed methods research has emerged within the past decade as a "value-added" research methodology that can add depth of analysis and explanation to conventional research, thus offering "the best of both" forms of research by drawing on the respective strengths of qualitative and quantitative forms of research. Such research designs, when crafted with high rigor, can provide more accurate and useful results that offer greater depth of information and explanation as these can inform the design of more efficacious prevention interventions.

Acknowledgments This chapter was supported by grant number P20 MD 002207-05 and 5P20MD002287 from the National Institute on Minority Health and Health Disparities (NIMHD), Elias Provencio-Vasquez, Principal Investigator. The content is solely the responsibility of the authors and does not necessarily represent the official view of the NIMHD, the National Institutes of Health, or the National Institute of Child Health and Human Development (NICHD). Funding was also provided to the third author by an institutional training grant from the Nation Institute of Mental Health (T32 MH018387) and by a National Research Service Award from NICHD (1F31 HD070602).

References[1]

*Armstrong, M. I., & Boothroyd, R. A. (2007). Predictors of emotional well-being in at-risk adolescent girls: Developing preventive intervention strategies. *Journal of Behavioral Health Services Research, 34*, 435–453.

Biglan, A., Domitrovich, C., Ernst, J., Etz, K., Mason, M. J., Robertson, E., et al. (2011). *Standards of knowledge for the science of prevention*. Fairfax, VA: Society for Prevention Research.

*Bowleg, L., Teti, M., Malebranche, D. J., & Tschann, J. M. (2012). "It's an uphill battle everyday": Intersectionality, low-income Black heterosexual men, and implications for HIV prevention research and interventions. *Psychology of Men and Masculinity*. doi:10.1037/a0028392. Advance online publication.

Brown, R. C., Colditz, G. A., & Proctor, E. K. (2012). *Dissemination and implementation research in health: Translating science to practice*. New York: Oxford University Press.

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research, 1*, 8–22.

*Cappella, E., Jackson, D. R., Bilal, C., Hamre, B. K., & Soule, C. (2011). Bridging mental health and education in urban elementary schools: Participatory research to inform intervention development. *School Psychology Review, 40*, 486–508.

Castro, F. G., & Coe, K. (2007). Traditions and alcohol use: A mixed methods analysis. *Cultural Diversity and Ethnic Minority Psychology, 13*, 269–284.

Castro, F. G., Kellison, J. G., Boyd, S., & Kopak, A. (2010). A methodology for conducting integrative mixed-methods research and data analyses. *Journal of Mixed Methods Research, 4*, 342–360.

Center for Scientific Review. (2013, January). Definitions of criteria and considerations for research project grant (RPG/R01/R03/R15/R21/R34) critiques. Retrieved from http://grants.nih.gov/grants/peer/critiques/rpg.htm.

*Chakrapani, V., Newman, P. A., Shunmugam, M., & Dubrow, R. (2011). Socio-structural contexts of needle and syringe sharing behaviours of HIV-positive injecting drug users in Manipur, India: A mixed methods investigation. *Harm Reduction Journal, 8*, 1–10.

*Cianelli, R., Ferrer, L., & McElmurry, B. J. (2008). HIV prevention and low-income Chilean women: Machismo, marianismo and HIV misconceptions. *Culture, Health and Sexuality, 10*, 297–306.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

*Coombes, L., Allen, D., Marsh, M., & Foxcroft, D. (2009). The Strengthening Families programme (SFP) 10–14 and substance misuse in Barnsley: The perspectives of facilitators and families. *Child Abuse Review, 18*, 41–59.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011). *Best practices for mixed methods research in the health sciences*. Bethesda, MD: Office of Behavioral and Social Science Research.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advances in mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Thousand Oaks, CA: Sage.

Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research, 6*, 80–88.

Denzin, N. K., & Lincoln, Y. S. (1994). Entering the field of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 1–17). Thousand Oaks, CA: Sage.

Diez Roux, A. V. (2011). Complex systems thinking and current impasses in health disparities research. *American Journal of Public Health, 101*, 1627–1634.

[1] Studies identified by an asterisk (*) are reviewed in Table 20.1.

Dreher, M. (1994). Qualitative research methods from the reviewer's perspective. In J. M. Morse (Ed.), *Critical issues in qualitative methods* (pp. 281–297). Thousand Oaks, CA: Sage.

Fielding, N. G. (2012). Triangulation and mixed methods designs: Data integration with new research technologies. *Journal of Mixed Methods Research, 6*, 124–136.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfriedson, D., Kellam, S., et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*, 151–175.

Gelo, O., Braakman, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative Psychological and Behavioral Science, 42*, 266–290.

Gonzalez-Guarda, R. M., Florom-Smith, A. L., & Thomas, T. (2011). A syndemic model of substance abuse, intimate partner violence, HIV infection, and mental health among Hispanics. *Public Health Nursing, 28*, 366–378.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Thousand Oaks, CA: Sage.

Hanson, W. E., Creswell, J. W., Plano Clark, V. L. P., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology, 52*, 224–235.

Hayes, A. F. (2013). *An introduction to mediation, moderation and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory and practice*. New York: Guilford Press.

Hesse-Biber, S. (2012). Feminist approaches to triangulation: Uncovering subjugated knowledge and fostering social change in mixed methods research. *Journal of Mixed Methods Research, 6*, 137–146.

Homer, J. B., & Hirsch, G. B. (2006). System dynamics modeling for public health: Background and opportunities. *American Journal of Public Health, 96*, 452–458.

*Hanbury, A., Wallace, L. M., & Clark, M. (2011). Multiple outcome measures and mixed methods for evaluating the effectiveness of theory-based behavior-change interventions: A case study targeting health professionals' adoption of a national suicide prevention guide. *Psychology, Health and Medicine, 16*, 291–303.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*, 112–133.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American Journal of Occupational Therapy, 45*, 214–222.

Landsverk, J., Brown, C. H., Chamberlain, L., Palinkas, L., Ogihara, M., Czaja, S., et al. (2012). Design and analysis of dissemination and implementation research. In R. C. Brownson, G. A. Colditz, & E. K. Protcor (Eds.), *Dissemination and implementation research in health: Translating science to practice* (pp. 225–260). New York: Oxford University Press.

Lassiter, P. S., Dew, B. J., Newton, K., Hays, D. G., & Yarbrough, B. (2006). Self-defined empowerment for gay and lesbian parents: A qualitative examination. *The Family Journal: Counseling and Therapy for Couples and Families, 14*, 245–252.

*Lewis, I., Maruia, B., & Walker, S. (2008). Violence against women in Papua New Guinea. *Journal of Family Studies, 14*, 183–197.

Little, T. D., & Freeman, M. (2011). Why the society on qualitative inquiry would be good for division 5. *Score Newsletter, 33*, 6.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84–99.

Mertens, D. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 135–164). Thousand Oaks, CA: Sage.

Mertens, D. M. (2012). Triangulation and mixed methods research: Provocative positions. *Journal of Mixed Methods Research, 6*, 75–79.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Milinki, A. K. (1999). *Cases in qualitative research: Research reports for discussion and evaluation*. Los Angeles, CA: Pyrczak Publishing.

*Montaghan, M., Sanders, R. E., Kelly, K. P., Cogen, F. R., & Streisand, R. (2011). Using qualitative methods to guide clinical trial design: Parent recommendations for intervention modification in type 1 diabetes. *Journal of Family Psychology, 25*, 868–872.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*, 120–123.

*Nelson, K., & Tom, N. (2011). Evaluation of substance abuse, HIV and hepatitis prevention initiative for urban Native Americans: The Native Voices Program. *Journal of Psychoactive Drugs, 43*, 349–354.

*O'Brien, R. A., Moritz, P., Luckey, D. W., McClatchey, M. W., Ingoldsby, E. M., & Olds, D. L. (2012). Mixed methods analysis of participant attrition in the Nurse-Family Partnership. *Prevention Science, 13*, 219–228.

Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & R. C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 351–383). Thousand Oaks, CA: Sage.

Patten, M. L. (2012). *Understanding research methods: An overview of the essentials* (8th ed.). Glendale, CA: Pyrczak.

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*, 209–233.

Sharp, J. L., Mobley, C., Hammond, C., Withington, C., Drew, S., Stringfield, S., et al. (2012). A mixed methods sampling methodology for a multisite case study. *Journal of Mixed Methods Research, 6*, 34–54.

Siebers, P. O., Macal, C. M., Garnett, J., Buxton, D., & Pidd, M. (2010). Discrete-event simulation is dead, long live agent-based simulation! *Journal of Simulation, 4*, 204–210.

*Simoes, C., Matos, M. G., & Batista-Foguet, J. M. (2008). Juvenile delinquency: Analysis of risk and protective factors using quantitative and qualitative methods. *Cognitive Brain Behavior, 12*, 389–408.

Sue, D. W., & Sue, D. (1999). *Counseling the culturally different: Theory and practice* (3rd ed.). New York: Wiley.

Tang, W., & Bennett, D. A. (2010). The explicit representation of context in agent-based models of complex adaptive spatial systems. *Annals of the Association of American Geographers, 100*, 1128–1155.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed-methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Sage handbook of mixed methods in social and behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.

Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 3–50). Thousand Oaks, CA: Sage.

Triandis, H. C., McCusker, C., & Hui, C. H. (1990). Multimethod probes of individualism and collectivism. *Journal of Personality and Social Psychology, 59*, 1006–1020.

*Tummala-Narra, P., Inman, A. G., & Ettigi, S. P. (2011). Asian Indians' responses to discrimination: A mixed-method examination of identify, coping, and self-efficacy. *Asian American Journal of Psychology, 2*, 205–218.

*Wilkerson, J. M., Danilenko, G. P., Smolenski, D. J., Myer, B. B., & Rosser, B. R. S. (2011). The role of critical self-reflection of assumptions in an online HIV intervention for men who have sex with men. *AIDS Education and Prevention, 23*, 13–24.

Yoshikawa, H., Weisner, T. S., Kalil, A., & Way, N. (2008). Mixing qualitative and quantitative research in developmental science: Uses and methodological choices. *Developmental Psychology, 44*, 344–354.

Part IV
Analytic Methods

Chapter 21
Latent Class Analysis in Prevention Science

Karen Nylund-Gibson and Shelley R. Hart

Introduction

Often prevention researchers have an innate sense that important and noticable difference exist in the popoulatoin that they are studying. In fact, it is usually of utmost importance to identify and understand the meaningful, yet often subtle, differences that might exist in a population. Latent class analysis (LCA) has a rich history as an analytical tool for many disciplines and provides a unique opportunity for prevention scientists to better understand those nuanced differences—that is, heterogeneity—in a population. A better understanding of the population is likely to lead to better treatment outcomes, such as more specifically tailored interventions. Further, the intervention or program may have different effects for different parts of a targeted population; thus, a one-size-fits-all approach to studying intervention effects will not properly describe the ways in which the intervention made a difference. In this way, LCA offers the opportunity to allow for treatment effects to be estimated separately for different subgroups of interest—that is, to investigate potential mechanisms by which certain effects are produced.

A variety of questions of interest for prevention researchers might lend themselves to the LCA methodology such as: What are the factors that place an individual at risk for developing a particular health challenge (e.g., depression, diabetes)? Can we tailor our interventions to best benefit different people? What protective factors might exist in the environment or the individual (e.g., parent monitoring, temperament) that mitigate the impact of those risk factors? How is it that individuals with similar levels of risk ultimately demonstrate different outcomes? How do individuals with different experiences have similar outcomes? Why does an intervention work for one person but not another? Thus, familiarity with this type of latent variable modeling may be of particular utility to prevention scientists.

K. Nylund-Gibson (✉)
Department of Education, Gevirtz Graduate School of Education, University of California, Santa Barbara, CA, 93106-9490, USA
e-mail: knylund@education.ucsb.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_21,

This chapter offers a nontechnical introduction to the LCA model. We focus on LCA models with binary items (or indicators), though it should be noted that LCA models with continuous items, called latent profile analysis (LPA), are possible but out of the scope of this chapter. We highlight important model parameters and their utility in prevention science and discuss some recent applications of LCA and their utility in understanding heterogeneity. To illustrate the model, we present a data example using LCA to study classes of adolescents' relationships with adults outside the family and link these to future behaviors of interest. Lastly, modeling extensions and future modeling developments are discussed.

LCA as a Way to Describe Population Variation

The term *modeling heterogeneity* is often used to describe a statistical modeling approach that aims to uncover and describe the qualitative differences in a population. Understanding the heterogeneity among individuals within a targeted population provides the opportunity to answer many of the questions, such as those above, which are asked by researchers and, ultimately, to design better treatment measures and intervention efforts. Stated another way, uncovering the way individuals are similar (i.e., examining how individuals grouped by certain characteristics differ from those without certain characteristics) provides a crucial foundation for understanding outcomes in the social sciences.

In general, sources of heterogeneity can be directly observable (e.g., gender, race, academic grade level) or unobserved. An example of unobserved heterogeneity would be the phenomenon that not everyone living in a risky neighborhood will develop behavior risk profiles. Another example would be that not everyone involved in a targeted intervention responds equally to treatment. There may not be a single reason for these outcomes, but some combination of variables may lead to this unobserved heterogeneity among individuals.

A wide range of models, called finite mixture or simply mixture models (e.g., LCA), aim to illuminate the unobserved heterogeneity in a population of interest. The fundamental idea in mixture modeling is that there are substantively interesting and important variations within a larger population and that identifying and understanding that heterogeneity provides a more accurate and nuanced view of the population. The goal of the analysis is to identify and enumerate this heterogeneity by creating groups of individuals that have similar profiles but are different from other groups in substantively qualitative ways. Prevention science researchers have given more attention to models that focus on capturing these individual differences (e.g., Collins & Lanza, 2010; Logan, Hall, & Karch, 2011; Muthén, Jo, & Brown, 2003).

Muthén's (2001) latent variable modeling framework provides a good lens to organize these models in respect to other more commonly known models (see Table 21.1). Using this framework, models can be classified using two criteria: (1) the scale of the latent variable (i.e., continuous, categorical, or hybrid) and

Table 21.1 Muthén's (2001) latent variable models

	Continuous latent variables	Categorical latent variables	Hybrids
Cross-sectional models	Factor analysis	Regression mixture analysis	Factor mixture analysis
	Structural equation models	Latent class analysis	
Longitudinal models	Growth analysis (random effects)	Latent transition analysis	Growth mixture analysis
		Latent class growth analysis	

(2) the type of data (i.e., cross-sectional or longitudinal). The more traditional latent variable models, factor analysis (FA) and structural equation modeling (SEM), use cross-sectional data and continuous latent variables. Growth analysis is an extension of these two methods to longitudinal data by using latent factors as random effects, where the change in individuals' scores over time is modeled. Four additional models—that is, regression mixture analysis, LCA, latent transition analysis (LTA), and latent class growth analysis (LCGA)—use categorical latent variables. The first two types are cross-sectional models, whereas the last two are longitudinal models. Finally, hybrids of the aforementioned models include newer modeling techniques (i.e., factor mixture analysis and growth mixture analysis) that allow for both the classification of individuals into latent classes and the calculation of continuous latent scores, each associated with a selected class.

What Is Latent Class Analysis?

LCA is a model-based approach to creating group of individuals, similar to cluster analysis. Unlike other group comparison approaches, in LCA group membership is not directly observed. In a multiple group approach, a variable such as gender is used to group individuals, and then modeling results are compared across males and females. As mentioned, in LCA group membership is not directly observed; rather it is latent and is derived based on the set of variables selected by the researchers.

This approach to creating groups can be used as an alternative to using cut scores or mean- or median-split groups. The essence of LCA is that it uses individual responses to a *set* of variables as a way to create groups. Because of the use of multiple measured items to create the groups, the clustering of individuals that emerge from LCA will vary from those when using a mean or median split. Often, the groups that emerge from LCA may not be patterns that would have been otherwise identified, since looking at items individually does not allow for the identification of the patterns that emerge based on the set of items.

In addition, LCA can be an alternative to creating a sum score. This commonly used method of summing up dichotomous items to create a composite score

provides a way to create a continuous score based on a set of items; however, there are several disadvantages to this approach. First, it assumes that all items on the measure are equally good at capturing the outcome of interest. This is commonly not the case since at times there are items we know are more important than others. Further, when using sum scores, there are multiple ways to achieve the same score.

Consider the following example. The Kindergarten Student Entrance Profile (KSEP; Lilles et al., 2009; Santa Maria–Bonita School District, First 5 of Santa Barbara County, & University of California Santa Barbara, 2005) is a measure intended to assess the readiness of children for kindergarten. In some applications, a sum score was created based on the count of skills that a child mastered among the 16 items spanning social-emotional, physical, and cognitive domains. Students with a score of 7 could have any constellation of items mastered—one student may have mastered all of the social-emotional items but no cognitive items and received a 7, while another student could have mastered all the cognitive items and one of the social-emotional items with a score of 7—but the items mastered could be entirely different. Teachers could easily identify students at the extremes—students who are either clearly ready or not ready for kindergarten. However, the students in the middle are more difficult to classify. LCA on the 16 KSEP items revealed two extreme classes (i.e., a high and low class as would be expected) as well as four other classes that were more in the middle (Quirk, Nylund-Gibson, & Furlong, 2013). These middle classes, while differing on the set of items endorsed, would likely have resulted in similar sum scores. Important differences emerged from these classes. When investigating second grade reading achievement, the latent classes that had mastered the cognitive items were more likely to have higher achievement. That is, not all of the "middle" classes—students who would have had similar sum scores—did equally well on later test scores. Thus, LCA provided a way to understand the heterogeneity in more detail than sum scores could.

Applications of LCA in Prevention Research

The example above illustrates one way LCA can provide more insight into understanding population heterogeneity over traditional methods. Before discussing the LCA model and specific considerations of model fitting and selection, it is useful to highlight a few more applications of LCA in prevention science, demonstrating the utility and insight gained from the application of this model. Typically, the use of LCA involves the discovery of the latent class structure of a population, which is then linked with covariates or outcome variables. For example, to evaluate the outcomes of academic and behavior problems in school, Reinke, Herman, Petras, and Ialongo (2008) identified subgroups of children based on measures collected in the fall of first grade, including teacher ratings of conduct and attention problems and student-assessed verbal and quantitative skills. Due to suspected gender differences, LCA models were fit separately for male and female students allowing for

the evaluation of different class solutions by gender. For boys, four classes emerged that classified these male students: (1) those with no problems, (2) those with behavior problems only, (3) those with academic problems only, and (4) those with both academic and behavior problems. For girls, three classes emerged, interpreted similarly, but with the absence of a class of students exhibiting only behavior problems. Thus, by allowing for differences in the measurement of the latent classes by gender, a known source of heterogeneity, different classes emerged.

Next, Reinke et al. (2008) used these classes to predict a variety of distal outcome variables that were measured in sixth grade using logistic regression including deviant peer affiliation, conduct problems, poor grades, presence of special education, and suspensions from school. By estimating the classes by gender, it was possible to examine differences in outcomes for boys and girls. For example, results indicated that boys with only academic problems did not significantly differ from the reference class (no problems), whereas girls demonstrated significantly worse grades and a higher likelihood for receiving special education services. In this way, the use of the covariate elucidated important differences that might have not been notable otherwise. Findings such as these have important implications for prevention science as interventions can be more specifically tailored to drive effectiveness.

Another use of LCA in a research paradigm is to use the latent class variable as an outcome variable. For example, to compare two assessment tools to identify dysregulated children, Althoff, Ayer, Rettew, and Hudziak (2010) employed LCA to identify profiles of children based on several subscales of the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) recently indicated in the assessment of youth with bipolar disorder. These researchers posited that this dysregulation profile (CBCL-DP) would identify similarly dysregulated children (i.e., high on attention problems, aggression, and anxiety depression) as another subscale attempting to capture traumatized children (CBCL-PTSP). Forty-one items from the CBCL were employed as indicators for the LCA. Several LCA models were analyzed, with both covariates included and not included. Fit indices were investigated, and a seven-class model was chosen with gender as a covariate. Two of these classes were selected as outcome variables due to their dysregulation profiles and investigated through subsequent Receiver Operating Characteristic Curve (ROC) analyses. That is, individuals were assigned to the class they most likely belonged to using posterior probabilities, and the CBCL-PTSP scores were used as the classifier to predict class membership. In this way, the latent class membership was the outcome of interest, and the ability to predict class membership was the focus of the study.

Other applications of LCA have also been used recently in more unique ways, including using LCA as a confirmatory method (see Finch & Bronk, 2011), using LCA to identify and describe change patterns in repeated measure data (see Lanza & Collins, 2006), and using LCA in a traditional MIMIC model context (see Yang, 2005).

Basic Elements of LCA

Research Questions

As the goal of LCA is to identify and model heterogeneity in the population, research questions are driven by attempts to understand this heterogeneity. LCA is an exploratory analysis, and as such the research questions are focused on the following questions: (1) Are there distinct heterogeneous latent classes in the sample population? (2) If so, how many latent classes are there? (3) Are there differences in membership across these groups with respect to other variables (covariates and distal outcomes)?

Researchers decide to do a clustering approach like LCA because they believe that there are subgroups in the population but cannot identify them, so question 1 is commonly supported. Research question 2 involves a complex class enumeration process described below that is often the most time-consuming part of the modeling process. And lastly, once the number of classes is decided on, question 3 addresses how the classes relate to other variables that are hypothesized to be related to the emergent latent classes.

Choice of Items

Directly related to the research question is the choice of items used in the analysis. The groups that emerge in LCA are determined by the measured items used in the analysis. Therefore, a critical step in LCA modeling is the selection of items. This process should be guided by strong substantive reasoning. Since the variation in the selected items is used to define the latent class variable, it is critical that a thoughtful selection is used that will define the heterogeneity believed to exist. Further, the items should capture characteristics that are important and relevant while being able to distinguish among the resulting classes.

It is possible and not surprising that an LCA model that uses, say, seven items would result in a different solution with respect to the number and type of classes that emerge compared with a model that is fit with the same seven items plus three more. This is because the classes describe the heterogeneity in the set of items used in the analysis and the emergent latent classes best separate individuals from one another such that they are similar in their responses to those within a given class but different from others in other classes. If items are describing the heterogeneity among the set of items, it then follows that if you change the items, even by just adding three more, the emergent classes may change.

LCA Model Parameters

The two parameters of interest in an LCA model are the conditional item-response probabilities and class probabilities. Similar to factor loadings in factor analytic models, these two parameters are used to create substantive labels for the emergent latent classes. Item-response probabilities describe an individual's probability of endorsing an item, given the class (or group) they belong to (i.e., conditional on the class membership). These probabilities—particularly the *pattern* or *profile* of probable responses for a given individual in each class—provide important information in understanding and assigning labels to the classes (Collins & Lanza, 2010). The class probabilities describe the relative size of the classes.

The conditional item probabilities are used to attach substantive meanings to each class, much like factor scores are used to understand and label the factors in factor analysis. These values are plotted in an item probability plot to aid in the interpretation of the latent classes. Figure 21.1 displays two different LCA item probability plots using five observed items. Along the x axis are the five observed items, and on the y axis are the conditional item probabilities for each class.

The other key parameter of interest in LCA is the class probability. This parameter describes the relative size of the emergent latent classes. These parameters are also used for describing and understanding the classes because they provide a sense of whether a class is a normative class (e.g., a large class probability) or if it represents a more rare profile (e.g., small class probability). In Fig. 21.1, the classes are approximately equally probable as described in the legend (i.e., Class 1 includes 41 % of the sample and Class 2 includes 59 %).

Ordered Versus Unordered LCA Solutions

The top panel of Fig. 21.1 is an example of an *ordered solution,* and the bottom panel is an example of an *unordered solution*. The ordered solution panel in Fig. 21.1 displays two profiles that do not cross. The top profile (Class 1) is characterized by having a relatively high probability of endorsing all the items, whereas the lower profile (Class 2) is characterized by having a low probability of endorsing all the items. This is considered an ordered solution because the profiles do not cross. For example, if the two-class solution on the top panel had items that measured violence exposures, Class 1 would be interpreted as the "high-exposure" class and Class 2 as the "low-exposure" class.

In practice, ordered solutions are more likely to occur when the observed items are all measuring the same construct (e.g., five items of depression) where the emergent classes capture differing levels of the underlying construct. One may

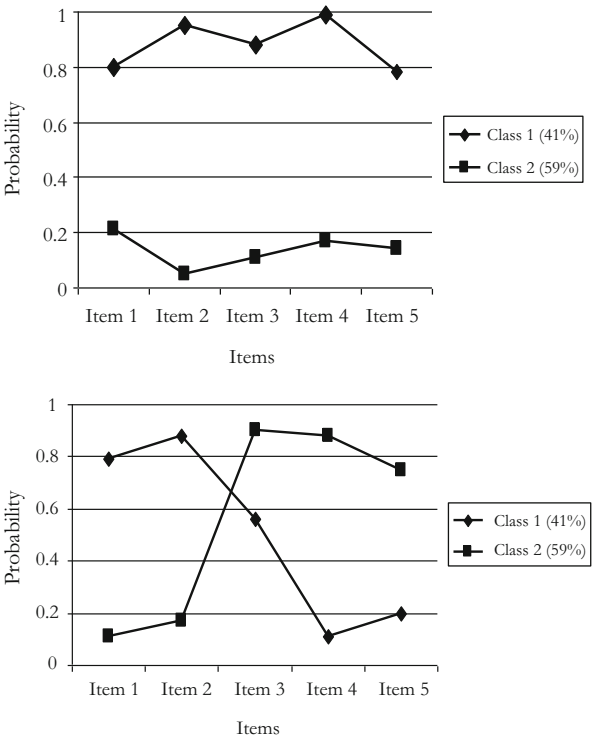

Fig. 21.1 Item probability plots for ordered (*top*) and unordered (*bottom*) LCA solution with binary outcomes

argue in this context that latent classes are not needed with ordered classes since the single dimensional construct could be captured as a latent factor. However, LCA may serve as a non-normal type of factor since the distribution of the classes (e.g., class sizes) may not be normally distributed. Further, it may be of substantive interest to create groups of people with different levels of the latent construct, and factor analysis does not do this. Thus, even with an ordered solution, LCA provides insight into heterogeneity that factor analysis may not.

The bottom panel in Fig. 21.1 is an unordered latent class solution where the class profiles do in fact cross each other. The profiles in the bottom panel of Fig. 21.1 show that items 1, 2, 4, and 5 are significant in differentiating the two profiles. Consider again an example of violence exposure. If items 1 and 2 in Fig. 21.1 were community violence items, Class 1 would be interpreted as a "community violence exposure" class because the class is characterized by having a high probability of endorsing the first two items, but not the other items. If items 4 and 5 measured exposure to peer violence, Class 2 would be the "peer violence exposure" class because individuals in this class have a high probability of endorsing the peer violence items, but not the rest of the items.

There is a wide range of possibilities of what the classes look like in an unordered solution, and the possibilities increase when the number of classes increases. For example, it is possible to have one or two ordered classes among a

few other unordered classes, such as described earlier in the Quirk et al. (2013) KSEP example. Unordered classes present another advantage of LCA—these groups of individuals would not likely be easily identifiable by an expert rater since it is the variation on the combination of items that is driving the differences in an unordered solution. Thus, LCA provides a nuanced way to understand the heterogeneity that exists in the population that other more traditional methods may not be able to capture.

LCA in Practice

Model Fit

Beyond the selection of measured items to use, the next critical step in LCA models involves deciding on the number of latent classes, or groups, that are necessary. In practice, when estimating LCA models, researchers fit a series of models starting with one class and increasing the number of classes until non-meaningful solutions are achieved or the model results in nonconvergence. The decision regarding the number of classes is an important step in the modeling process and often involves the consideration of several models and their substantive interpretation. Substantive theory, classification quality, and the fit of the model to the data are factors that drive the decision for which model to evaluate. An array of model fit statistics exist, none of which are perfectly reliable (Nylund, Asparaouhov, & Muthén, 2007).

As a result of the imperfect nature of any one fit statistic, an unresolved issue in the application of LCA models (or any type of mixture model for that matter) is how best to decide on the number of classes, and as a result, several different indices are used. The Bayesian information criterion (BIC) is a trusted fit index, where the model with the lowest BIC is preferred. In addition, other fit statistics like the bootstrap likelihood ratio test (BLRT), Lo-Mendell-Rubin (LMR) adjusted likelihood ratio test, and quasi-Bayesian information-heuristic model fit comparisons are more recently being used to evaluate fit of LCA models (for more, see Masyn, 2013; Nylund et al., 2007).

In addition to fit statistics, the plausibility of the model is also studied. That is, the statistical tools are useful in helping identify a model that provides the best statistical fit, but if the results do not make substantive sense, then it calls into question how trustworthy and useful the results are. The congruence of one or two statistical fit indicators (e.g., the BIC and LMR), along with a plausible and interpretable solution, is often considered adequate justification to trust the results of an LCA analysis. Finally, once the number of classes is established, covariates and distal outcomes are added to that model to explore the characteristics and outcomes of the people in each class.

Software

Several software packages can be used to estimate LCA models. Among the most common is Mplus V7 (Muthén & Muthén, 1998/2012), a general statistical software package that is used in the empirical example in this chapter. Mplus can be used to estimate mixture models, structural equation models, and a wide range of other latent variable models. Mplus leads the way in terms of the integration of mixture models with other types of latent variable models including multilevel mixture modeling, factor mixture models, and longitudinal mixture models, to name a few. Other software packages that can be used to estimate LCA models include Latent GOLD V4.5 (Statistical Innovations, Inc., 2005–2011), PROC LCA that is used in SAS (Lanza, Dziak, Huang, Xu, & Collins, 2011), and GLLAMM for Stata (Rabe-Hesketh, Skrondal, & Pickles, 2004).

Covariates and Distal Outcomes

Once the classes are enumerated, attention is often turned to understanding who comprises the latent classes and if there are any consequences of membership in these classes. This involves including covariates and distal outcomes in the analysis. For covariates, which can be either categorical or continuous, this involves a multinomial logistic regression of the latent class variable onto the covariates. As in traditional logistic regression, this results in a comparison of the odds of being in the latent classes compared with a reference latent class for different values of the covariate. It is possible to change the reference class so that all of the emergent latent classes can be compared with respect to the covariate, but most traditionally, one meaningful reference class is identified (e.g., the largest, the normative, the most or least at risk) and all covariate comparisons are made with respect to that class.

With distal outcomes, we do not simply regress the distal outcome onto the latent class variable. Instead, we estimate the mean of the distal outcome for each latent class and then compare means across the classes to assess whether there is a difference on a distal outcome. Several different methods have been proposed for doing this and are available in the common statistical software package Mplus, but the most current method believed to produce the most unbiased results is the Three-Step Method (Asparouhov & Muthén, 2012; Vermunt, 2010). As the name suggests, this method is a three-step estimation method that accounts for the fact that the model allows individuals not to be perfectly assigned to class. That is, the estimation of the means for the distal outcomes for each class accounts for the non-perfect assignment.

Posterior Probabilities

Classifying individuals into one of their assigned classes may be of interest after estimating an LCA model. Similar to estimating factor scores, posterior class probabilities can be estimated.[1] These probabilities are an individual's probability of being in each latent class of the fitted model, given the individual's observed response pattern on the set of items. Model class assignment can be done whereby individuals are assigned to the latent class for which they have the highest posterior probability of membership. For example, consider a three-class LCA solution. An individual who has posterior probabilities of being in the three classes of 0.80, 0.15, and 0.05 for Classes 1, 2, and 3, respectively, would be assigned to Class 1 because that is where the highest posterior probability is observed.

Data Example: LCA to Study Patterns in Students' Relationships with Adults Outside Their Family

To illustrate the previous discussion of the LCA model, an example is provided. The purpose of the analysis was to examine the connections or relationships students have with adults outside their family and the effect those relationships have on later outcomes of interest (e.g., suicidal ideation, substance use). These outcomes were chosen as they may be a more proximal indicator of future challenges (e.g., major depressive disorder, drug dependence disorders). If these outcomes are shown to be associated with adult relationships, a potential avenue for intervention may be illuminated.

These data come from a study conducted in one school district in central California as part of the statewide California Healthy Kids Survey (CHKS; available from http://chks.wested.org/administer). CHKS is a series of assessment modules administered to students in schools across the state to better understand the resilience of students and the protective and risk factors associated with outcomes of interest (e.g., drug use). It was developed by WestEd in conjunction with Duerr Evaluation Resources for the California Department of Education, and substantial psychometric evidence is available (e.g., Hanson & Kim, 2007; Waters & Cross, 2010). Results from this survey have been used to inform policy at the state and local levels, and research resulting from administration of CHKS has promoted understanding of students' experiences in areas such as school connectedness, victimization, student engagement, risky behaviors, and physical health (e.g., Davis & Carpenter, 2009; Felix, Furlong, & Austin, 2009; Sharkey, You, & Schnoebelen, 2008). For the most part, research investigating these experiences has been cross-sectional.

[1] In Mplus, posterior probabilities can be saved to an external data file using the "SAVEDATA" command and specifying "save = cprobabilities."

Two sections of CHKS were used in the current study as the observed items for the LCAs. Twelve items were selected to investigate students' perceptions of their relationships with and support from both school-based adults (i.e., teachers and other staff) and adults found outside the school system (e.g., coaches, youth leaders). These items were almost identical with the exception of the stem (i.e., "At my school, there is a teacher or some other adult. . ." versus, "Outside of my home and school, there is an adult. . ."). These items included ". . .who really cares about me," ". . .who tells me I do a good job," ". . .who always wants me to do my best," and ". . .who believes that I will be a success." Two items were different: "At my school there is a teacher or adult who notices when I'm not there," "At my school there is a teacher or adult who listens to me when I have something to say," "An adult outside of the home or school notices when I am upset about something," and "I trust an adult outside of home or school."

The 12 items were originally measured on a 4-point Likert-type scale (e.g., "Not at all true" to "Very much true"). Due to distributional properties and ease of presentation for this example, the decision was made to dichotomize these items so that an endorsement of the item was considered if a student marked either "Pretty much true" or "Very much true." Students' responses to these items measured in eighth grade were used as the indicators for the LCA.

There has been research focusing on the feelings of connection a student has to his/her school, demonstrating the importance of this construct in relation to important outcomes, such as risky behaviors (e.g., substance use), intrapersonal violence (e.g., suicide ideation), and academic achievement (Chapman et al., 2010; Niehaus, Rudasill, & Rakes, 2012). School connectedness sometimes encompasses perceptions of support and connections with staff and teachers. Far less studied are the relationships and connections students have with individuals outside either the family or the school.

Several demographic covariates were used in the analyses, including age, gender, and ethnicity, to examine the impact of these characteristics on the patterns that emerge from the LCAs. The distal outcomes were both binary and continuous variables taken from the CHKS and several additional self-report measures administered to students. Distal outcomes were measured in tenth grade. The goal of the analysis was to examine the association of students' perception of their relationships with adults other than family members in middle school and various high school outcomes of interest. Briefly, three general areas were of interest including mood (i.e., suicidal ideation, optimism, life satisfaction), risky experiences (i.e., substance use, victimization, behavioral or emotional problems), school-related perceptions and behaviors (i.e., meaningful school participation, peer connectedness), and family connectedness.

Results

A cohort of $N = 157$ students completed several self-report measures at three different waves (fall/spring 2009, fall/spring 2010, fall/spring 2011).[2] At Wave 1, students were in eighth grade with an average age of 13.3 years (standard deviation of 0.54 years). The sample was primarily Hispanic (44.1 %) or White (42.2 %) with 54 % being girls and 43 % boys. Missing data for the items used in the LCA was minimal, with all variables missing less than 2 % of individuals' responses. Missingness on the covariates was also minimal (under 2 %); however, approximately 21 % of respondents were missing data on the distal outcome variables, measured in tenth grade. Examination of these individuals using Pearson's chi square and independent samples *t* tests indicated that they were not significantly different than those who had complete data on any covariates or items used for the LCA ($p \geq 0.05$). As full information maximum likelihood (FIML) was used, which has been demonstrated to be an appropriate way to deal with missingness when data are missing at random (Schafer & Graham, 2002), nothing further was done to address the missing.

Step 1: Class Enumeration

A total of four classes were estimated in the class enumeration process. The fit statistics for these models are displayed in Table 21.2. A nonpositive definite matrix was demonstrated in a four-class model, and the BLRT and the VLMR stay significant throughout estimation of the models. A scree plot of the BIC (see Fig. 21.2) reveals that values continue decreasing through a three-class model, initial inspection of a three-class model indicated potentially interesting results, and class size was still adequate. Therefore, a three-class model was selected.

Step 2: Addition of Covariates and Description of Classes

After the number of classes is determined, the next step is the prediction of class membership through the addition of covariates. For the current example, three covariates were added: age, gender, and race. To understand the relationship of the covariates to the classes, however, the item probability plots need to be examined and class names assigned. For Fig. 21.3, the probability of endorsing the item (i.e., pretty much or very much true) is graphed. As indicated by the figure,

[2] The authors would like to thank Drs. Erin Dowdy, Jill Sharkey, Erika Felix, and Mike Furlong for access to these data.

Table 21.2 Fit statistics for LCA of nonfamilial adult connectedness and support

# of classes	Log likelihood	BIC	ABIC	BLRT *p*-value	VLMR *p*-value	Entropy	Smallest class size
1	−938.74	1,938.08	1,901.48	–	–	1.00	157 (1.00)
2	−721.32	1,568.89	1,492.65	0.00	0.00	0.92	45 (0.29)
3	−645.91	1,483.72	1,367.82	0.00	0.00	0.93	19 (0.12)
4	Nonpositive definite						

Note: BIC Bayesian information criterion; *ABIC* adjusted Bayesian information criterion; *BLRT* bootstrap likelihood ratio test; *VLMR* Vuong-Lo-Mendell-Rubin likelihood ratio test

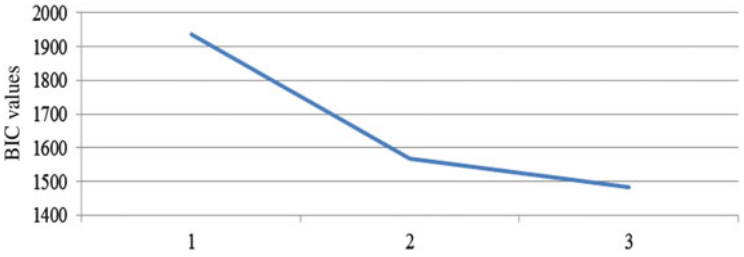

Fig. 21.2 Scree plot of the BIC for the class enumeration process

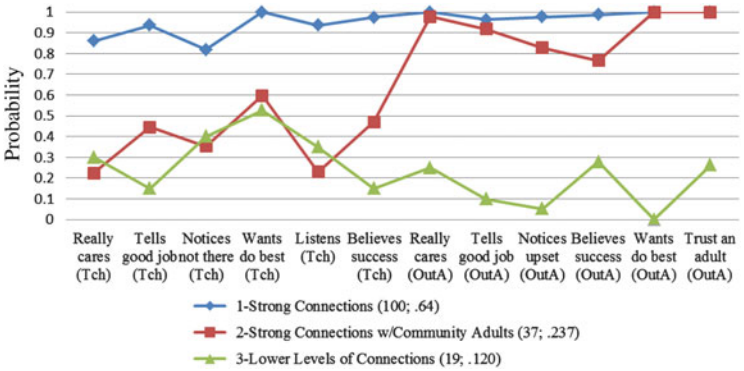

Fig. 21.3 Item probability plots for nonfamilial adult connectedness and support LCA *Note:* Probability of endorsing item is graphed on the *x*-axis. Classes are named and the number of individuals in each class is followed by the proportion of the sample (e.g., 100; 0.64 for Class 1 strong connections)

the three classes seem to be distinct, with the first class being the largest (0.64 of the sample) and most connected to adults outside the family (i.e., high rates of endorsement for all items). We named this class strong connections (SC). The second largest class (0.24) demonstrates fewer connections with adults at school, but seems to be strongly connected to adults outside school, and we named this class strong

connections with community adults (SCCA). Finally, the smallest class (0.12) appears to demonstrate the lowest levels of connections with adults outside the family, and we named this class lower levels of connections (LLC).

As the SC class appears to be the normative class (i.e., the largest and the one we might be the least concerned with in terms of risk), we discuss covariates using this as our reference class. Therefore, in comparison with this class, those in the SCCA class were three times more likely to be male ($p \leq 0.01$). No covariates significantly predicted membership in the LLC class.

Step 3: Investigating Distal Outcomes

Finally, to investigate the relationship of middle school students' connections with adults outside the family and various high school experiences, distal outcomes were added to the analyses. Table 21.3 lists the chi square statistic and post hoc tests to investigate further differences among the classes for the variables of interest. Differences among those in the SC class in comparison with those in the SCCA class existed on levels of optimism, meaningful participation in school, connections with peers, and family connections (all $p \leq 0.05$). Results were in an expected direction, with those in the SC class demonstrating higher levels of optimism, more meaningful participation in school, and stronger connections with peers and family members. Differences among those in the SC class in comparison with those in the LLC class existed on meaningful participation in school and family connections. Additionally, while not significant in the overall chi square test ($p = 0.06$), significant differences did exist between the SC and LLC classes on levels of behavioral and emotional problems ($p \leq 0.05$) and is reported in brackets in the table.

LCA Modeling Extensions: Latent Transition Analysis

The LCA model discussed in this chapter and illustrated in the preceding data example is a cross-sectional model. That is, it is a model used to describe heterogeneity in a population based on a single time point. With longitudinal data, it is possible to model how individuals transition out of the emergent classes over time using LTA. For example, if the nonfamilial adult connectedness and support were measured at two time points, we could estimate the latent classes that emerge at later time points and then model who remains stable in their relationships over time and who transitions into different relationship classes.

LTA is a widely used and powerful longitudinal extension of the LCA model that allows for the modeling of change among groups or clusters. In addition, LTA models, which are an autoregressive model of change, are a useful tool for research questions that involve change with two times points or when the change process is believed to be nonlinear.

Table 21.3 Distal outcome results

Distal outcome	Chi square	Post hoc
Mood		
Suicidal ideation	0.28	
Optimism	6.72*	SC vs. SCCA (4.94*)
Life satisfaction	2.28	
Risky experiences		
Substance use	1.90	
Alcohol	1.93	
Victimization	2.25	
Behavioral/emotional problems	5.74	[LLC vs. SC (4.37*)]
School related		
Meaningful participation	11.87**	SC vs. SCCA (4.89*)
		SC vs. LLC (10.36***)
Peer connectedness	16.66***	SC vs. SCCA (12.85***)
Family connections	7.62*	SC vs. SCCA (4.54*)
		SC vs. LLC (4.98*)

Note: ***$p \leq 0.001$; **$p \leq 0.01$; *$p \leq 0.05$; *SC* strong connections class, *SCCA* strong connections with community adults class, *LLC* lower levels of connections class. Post hoc tests indicate significant class difference. The class with the highest mean is listed first with chi square and significance listed in the parentheses

Future Modeling Directions with LCA

The integration of latent class models into larger latent variable models has been an active area of research that has led to a wide range of insightful applications. A recent development combines the multilevel framework to estimate the parameters of the LCA model. A multilevel LCA model not only accounts for the clustered nature of observations—for example, in communities, schools, hospitals, or regions—but also allows for inferences to be made about the impact of the clustering on individual's responses. For example, perhaps there is a characteristic of a hospital that leads to more patients having a better response to treatments.

Another more recent area of LCA model development is the possibility of estimating LCA models in the Bayesian framework. Using this Bayesian framework, which incorporates additional information in the modeling process in the form of priors, allows for several important modeling extensions. This includes more confidence in LCA modeling with an overall small sample. In addition, when using a Bayesian framework, we can feel more confident in the model estimates when one of the latent classes is really small, since it is often the case that the small latent classes are of interest to researchers.

Conclusion

The utility of LCA for the types of questions prevention scientists ask is undeniable. This nontechnical chapter highlighted the utility of the LCA model and the process of modeling. We provided an applied example to highlight that process: the steps taken, decisions made, and interpretations arrived at. We concluded with future directions being taken using LCA. The chapter provides a basic introduction to interested researchers and encourages them to access the rich resources that are available for this important data analytic tool.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, & Families.

Althoff, R. R., Ayer, L. A., Rettew, D. C., & Hudziak, J. J. (2010). Assessment of dysregulated children using the Child Behavior Checklist: A receiver operating characteristic curve analysis. *Psychological Assessment, 22*, 609–617. doi:10.1037/a0019699.

Asparouhov, T., & Muthén, B. (2012). *Auxiliary variables in mixture modeling: A 3-step approach using Mplus* (Mplus web note: No. 15). Retrieved from http://www.statmodel.com/examples/webnotes/webnote15.pdf.

Chapman, R. L., Buckley, L., Sheehan, M. C., Shochet, I. M., & Romaniuk, M. (2010). The impact of school connectedness on violent behavior transport risk-taking behavior, and associated injuries in adolescence. *Journal of School Psychology, 49*, 399–410. doi:10.1016/j.jsp.2011.04.004.

Collins, L. M., & Cliff, N. (1990). Using the Longitudinal Guttman Simplex as a basis for measuring growth. *Psychological Bulletin, 108*, 128–134.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Hoboken, NJ: Wiley.

Collins, L. M., & Sayer, A. G. (Eds.). (2001). *New methods for the analysis of change*. Washington, DC: American Psychological Association.

Compton, D. L., Fuchs, D., Fuchs, L. S., Elleman, A. M., & Gilbert, J. K. (2008). Tracking children who fly below the radar: Latent transition modeling of students with late-emerging reading disability. *Learning and Individual Differences, 18*, 329–337. doi:10.1016/j.lindif.2008.04.003.

Connell, A., Bullock, B. M., Dishion, T. J., Shaw, D., Wilson, M., & Gardner, F. (2008). Family intervention effects on co-occurring early childhood behavioral and emotional problems: A latent transition analysis approach. *Journal of Abnormal Child Psychology, 36*, 1211–1225. doi:10.1007/s10802-008-9244-6.

Davis, B., & Carpenter, C. (2009). Proximity of fast-food restaurants to schools and adolescent obesity. *American Journal of Public Health, 99*, 505–510. doi:10.2105/AJPH.2008.137638.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, 39*, 1–38.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 8*, 430–457.

Felix, E. D., Furlong, M. J., & Austin, G. (2009). A cluster analytic investigation of school violence victimization among diverse students. *Journal of Interpersonal Violence, 24*, 1673–1695. doi:10.1177/0886260509331507.

Finch, W. H., & Bronk, K. C. (2011). Conducting confirmatory latent class analysis using Mplus. *Structural Equation Modeling, 18*, 132–151. doi:10.1080/10705511.2011.532732.

Hanson, T. L., & Kim, J. O. (2007). *Measuring resilience and youth development: The psychometric properties of the Healthy Kids Survey* (Issues & Answers Report, REL 2007–No. 034). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West.

Kandel, D. B., & Jessor, R. (2002). The gateway hypothesis revisited. In D. B. Kandel (Ed.), *Stages and pathways of drug involvement: Examining the gateway hypothesis* (pp. 365–372). New York: Cambridge University Press.

Lanza, S. T., & Collins, L. M. (2006). A mixture model of discontinuous development in heavy drinking from ages 18 to 30: The role of college enrollment. *Journal of Studies on Alcohol, 67*, 552–561.

Lanza, S. T., Dziak, J. J., Huang, L., Xu, S., & Collins, L. M. (2011). *PROC LCA & PROC LTA user's guide version 1.2.6*. University Park, PA: The Methodology Center, Penn State.

Lazarsfeld, P., & Henry, N. (1968). *Latent structure analysis*. New York: Houghton-Mifflin.

Lilles, E., Furlong, M., Quirk, M., Felix, E., Dominguez, K., & Anderson, M. (2009). Preliminary development of the Kindergarten Student Entrance Profile. *The California School Psychologist, 14*, 71–80.

Logan, J., Hall, J., & Karch, D. (2011). Suicide categories by patterns of known risk factors: A latent class analysis. *Archives of General Psychiatry, 68*, 935–941. doi:10.1001/archgenpsychiatry.2011.85\.

Masyn, K. (2013). Latent class analysis and finite mixture modeling. In T. D. Little (Ed.), *The Oxford handbook of quantitative methods in psychology* (Vol. 2, pp. 551–611). New York, NY: Oxford University Press.

Muthén, B. O. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates. http://pages.gseis.ucla.edu/faculty/muthen/articles/Article_086.pdf.

Muthén, B., Jo, B., & Brown, C. H. (2003). Assessment of treatment effects using latent variable modeling: Comments on the New York School of Choice Study. *Journal of the American Statistical Association, 98*, 311–314. http://pages.gseis.ucla.edu/faculty/muthen/articles/Article_086.pdf.

Muthén, L. K., & Muthén, B. O. (1998/2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Niehaus, K., Rudasill, K. M., & Rakes, C. R. (2012). A longitudinal study of school connectedness and academic outcomes across sixth grade. *Journal of School Psychology, 50*, 443–460. doi:10.1016/j.jsp.2012.03.002.

Nishina, A., Bellmore, A., Witkow, M. R., & Nylund-Gibson, K. (2010). Longitudinal consistency of adolescent ethnic identification across varying ethnic contexts. *Developmental Psychology, 46*, 1389–1401. doi:10.1037/a0020728.

Nylund, K., Asparaouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: An Interdisciplinary Journal, 14*, 535–569. doi:10.1080/10705510701575396.

Quirk, M., Nylund-Gibson, K., & Furlong, M. (2013). Exploring patterns of Latino/A children's school readiness at Kindergarten entry and their relations with grade 2 achievement. *Early Childhood Research Quarterly, 28*(2), 437–449.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). Generalized multilevel structural equation modelling. *Psychometrika, 69*, 167–190.

Reboussin, B. A., Reboussin, D. M., Liang, K. Y., & Anthony, J. C. (1998). Latent transition modeling of progression of health-risk behavior. *Multivariate Behavioral Research, 33*, 457–478. doi:10.1207/s15327906mbr3304_2.

Reinke, W. M., Herman, K. C., Petras, H., & Ialongo, N. S. (2008). Empirically derived subtypes of child academic and behavior problems: Co-occurrence and distal outcomes. *Journal of Abnormal Child Psychology, 36*, 759–770. doi:10.1007/s10802-007-9208-2.

Santa Maria–Bonita School District, First 5 of Santa Barbara County, & University of California Santa Barbara. (2005). *Kindergarten student entrance profile*. Santa Barbara, CA: University of California, Santa Barbara.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147–177.

Sharkey, J. D., You, S., & Schnoebelen, K. (2008). Relations among school assets, individual resilience, and student engagement for youth grouped by level of family functioning. *Psychology in the Schools, 45*, 402–418. doi:10.1002/pits.20305.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis, 18*, 450–469.

Walrath, C. M., Petras, H., Mandell, D. S., Stephens, R. L., Holden, E. W., & Leaf, P. J. (2004). Gender differences in patterns of risk factors among children receiving mental health services: Latent class analysis. *Journal of Behavioral Health Services and Research, 31*, 297–311. doi:10.1007/BF02287292.

Waters, S., & Cross, D. (2010). Measuring students' connectedness to school, teachers and family: Validation of three scales. *School Psychology Quarterly, 25*, 164–177. doi:10.1037/a0020942.

Witkiewitz, K. (2008). Lapses following alcohol treatment: Modeling the falls from the wagon. *Journal of Studies on Alcohol and Drugs, 69*, 603–613.

Witkiewitz, K., & Villarroel, N. A. (2009). Dynamic association between negative affect and alcohol lapses following alcohol treatment. *Journal of Counseling and Clinical Psychology, 77*(4), 633–644. doi:10.1037/a0015647.

Yang, C.-C. (2005). MIMIC latent class analysis model for alcoholic diagnosis. *Structural Equation Modeling: A Multidisciplinary Journal, 12*, 130–147. doi:10.1207/s15328007sem1201_7.

Chapter 22
Discrete-Time Survival Analysis in Prevention Science

Katherine E. Masyn

Introduction

In prevention science, research questions regarding specific life course events, such as drug use, school suspension, GED obtainment, job loss, and marriage, are often concerned with both the *if* and *when* of event occurrence. For example, it may be of interest to investigate not only what predicts whether an individual commits a criminal offense but what predicts the age at which first offense occurs and what predicts the time between offenses; what predicts whether and when an individual may relapse after treatment for addiction; what predicts the incidence and duration of a major depressive episodes, etc. Furthermore, the actual timing or duration of an event may itself directly influence long-term risk for subsequent outcomes and may function as a mediator or moderator of putative risk and protective factors as well a preventive interventions. For example, a drug-use prevention program may have an indirect effect on the risk of adulthood substance use disorders by delaying the onset of substance use in adolescence.

Survival analysis refers to the general set of statistical methods developed specifically to model the timing of events while accommodating situations for which event time is only observed or known for a portion of the total sample. This chapter concerns a subset of those methods that deals with discrete-time events. I begin with a motivating data example that will be used for methods illustrations throughout. This is followed by a brief overview of survival analysis, including an explanation of the distinction between discrete- and continuous-time event histories and a presentation of the key quantities in a survival analysis—the hazard and survival probabilities. I then provide a detailed discussion of how one can construct event history records from longitudinal panel data and how such event history records can then be used in statistical discrete-time survival analysis.

K.E. Masyn (✉)
Harvard Graduate School of Education Cambridge, Harvard University, Cambridge, MA 02138, USA
e-mail: katherine_masyn@gse.harvard.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_22,

The chapter concludes with a consideration of more advanced modeling extensions and opportunities for prevention science applications.

Data Example

The data example for this chapter involves the age of drinking onset. As alcohol consumption by adults in the United States is commonplace, the "if" of drinking onset is not nearly as variable as the "when." Age of alcohol use onset has been found to be a strong predictor lifetime alcohol dependence and abuse (e.g., Grant & Dawson, 1997; Hingson, Heeren, & Winter, 2006). Additionally, there is also evidence that beginning to drink at an early age is associated with more immediate problems in adolescence and emerging adulthood, such as compromised brain development, risky sexual behaviors, poor school performance, and illicit drug use (Morean, Corbin, & Fromme, 2012). The ability to model the timing of onset can enable prevention scientists to effectively evaluate the potentially modifiable risk of early onset and prevention efforts targeted toward early onset users and aimed at the reduction of alcohol use disorders over the life course.

The data come from the National Longitudinal Survey of Youth 1997 (NLSY97; Ohio State University. Center for Human Resource Research, 1997)—a prospective longitudinal cohort study that has followed a nationally representative sample of US residents beginning in 1997 when participants were ages 12–16. The data used for the example in this chapter are from the first seven waves of data collection. Panel sampling weights for longitudinal analysis are used in all analyses and the household identifier was used to account for non-independence of youths sampled from the same household. Study participants with at least one wave of alcohol use data and complete data on the selected covariates—gender, race, parent education, and urban/nonurban area residency—were included in the final analysis data set, resulting in a total of 8,280 cases. This chapter does not address how to handle missing data on exogenous predictor variables of event histories; however, the available missing data techniques are the same as those available for multivariate regression and structural equation modeling (see, for example, Enders, 2010). All data analysis was done in the Mplus, version 7 modeling software (Muthén & Muthén, 1998–2012).

Survival Analysis Overview

A Brief History of Survival Analysis

Survival analysis methodology has its roots in life table data—for centuries people have been, both formally and informally, collecting population data on birth rates,

morbidity rates, and mortality rates. Typically, with regards to the mortality rates, information would be collected about the age of death, cause of death, etc. It is not surprising that at some point people wanted to be able to make projections or predictions about life expectancies and also make comparisons across subpopulations. For example, how does the risk of death change across the lifespan? Are men or women more likely to live past the age of 80? These questions began to develop into more sophisticated medical research questions about human survival. Spurred by such questions, statisticians in the early- and mid-twentieth century set to work on ways to model such data. The most noteworthy advance in survival analysis was made by David R. Cox in 1972. Cox introduced a semi-parametric approach to estimating a hazard function—the function describing the instantaneous rate of event occurrence across time—with continuous covariates. Since the presentation of the Cox regression model, there have been numerous other parametric, semi-parametric, and nonparametric models put forth in the survival analysis literature as well as alternative estimators that collectively make up modern day survival analysis techniques (see, for example, Hosmer, Lemeshow, & May, 2008; Singer & Willett, 2003).

The modeling approach Cox developed and most of the subsequent extensions have been focused on continuous-time scales for event occurrence (as the timing of events such as death occur in continuous time and are often recorded with a reasonable degree of precision). However, in his seminal 1972 paper, Cox also suggested an approach for modeling events occurring or recorded in discrete-time intervals. It is this type of discrete event history data that is the focus of this chapter.

Discrete-Time and Continuous-Time Survival Data

Time scales for events can be crudely classified as either continuous or discrete and the methods applied to one type of time scale do not necessarily apply to the other, just as regression techniques for continuous outcome variables do not apply directly to categorical outcomes. Applications of continuous-time methods assume that the timing of events is known exactly or that the discrete intervals on which time is measured are sufficiently small that is reasonable to treat the observed times as occurring on a continuous-time scale, assuming no two individuals have identical event times. Discrete-time events may be of two natures (Allison, 1982): (1) An event may occur at any point in time but only an interval of time during which the event occurred is recorded, e.g., a student may drop-out of school on any particular day of the school year, but data may only be available for the grade level at which the drop-out occurred. This is sometimes referred to as *grouped-time* survival data, e.g., all the days of the school year and subsequent summer are grouped together to form a 12-month time interval. (2) An event may only occur at discrete points in time, e.g., retention at a certain grade level only occurs at the end of a school year. For both types of discrete-time event histories, continuous-time survival models are clearly inappropriate and it is necessary to apply discrete-time analysis methods.

Survival and Hazard Functions

For different types of outcomes, there are standard statistical models that involve only select features of the outcome distribution. For example, with a linear regression model, we usually estimate the conditional mean and variance of the outcome given a set of covariates. Discrete-time event histories have their own set of distribution-related quantities that are substantively and statistically most appropriate for describing and modeling time-to-event processes. The two most prominent are the survival function and the hazard function.

Consider data that reflect an underlying continuous-time process with event times grouped into J discrete intervals, $(t_1,t_2, \ldots,t_J)$. Let the given sample consist of n independent individuals, $i = 1, \ldots, n$, with corresponding survival times, T_i, where $T_i = t_j$ if individual i experiences the event during time period t_j. The most intuitive representation of the distribution of event times is the survival function, describing the probability of an individual "surviving" (i.e., being event-free) beyond a particular time interval, given by

$$P_{\mathrm{S}}(t_j) = \Pr(T > t_j). \tag{22.1}$$

For drinking onset using 1-year age intervals, $P_{\mathrm{S}}(t_{15})$ would be the probability that drinking onset occurs *after* age 15; or, alternatively, $1 - P_{\mathrm{S}}(t_{15})$ is the proportion of adolescents that have first used alcohol by age 15 (i.e., *before* age 16).

In statistical modeling of the distribution of event times, the most common representation is the hazard function, describing the probability of an individual experiencing the event in a particular time interval among those individuals still event-free at the beginning of the interval, given by

$$P_{\mathrm{h}}(t_j) = \Pr(T = t_j | T \geq t_j). \tag{22.2}$$

For drinking onset using 1-year age intervals, $P_{\mathrm{h}}(t_{15})$ is the probability that drinking onset occurs *at* age 15 among adolescents that have *not* had their first use of alcohol prior to age 15.

Examining plots of the hazard probabilities is quite useful in understanding how risk for an event changes over time and how those changes influence the corresponding survival probabilities. For example, the life times of many living creatures follow a bathtub-shaped or U-shaped hazard function: we are most vulnerable at the beginning and end of our lifespan. Relating the shape of the hazard to the survival function, when the hazard is zero, the survival function is constant; when the hazard is high, the survival function is decreasing quickly; when the hazard is low, the survival function is decreasing slowly. The shapes of the hazard function and survival function are useful to examine together as the survival function not only reflects the cumulative risk impact on the population in each time period, it quantifies the proportion of the population susceptible to the risk indicated by the hazard for each time period. The proportion at risk for the event in time

period j is equal to $P_S(t_{j-1})$. Mathematically, the relationship between the hazard and survival probabilities is given by:

$$P_S(t_j) = \left(1 - P_h(t_j)\right) \cdot P_S\left(t_{j-1}\right). \tag{22.3}$$

The probability of surviving beyond time period j is equal to the probability of surviving beyond time period $j - 1$ multiplied by the probability of not experiencing the event in time period j. Alternatively,

$$P_S(t_j) = \prod_{m=1}^{j} (1 - P_h(t_m)). \tag{22.4}$$

Thus, even if the hazard probabilities are being estimated in the statistical model, it is possible to compute estimated survival probabilities as well.

The reason the hazard function is the standard representation of the distribution of event times for statistical models, even though the survival function is perhaps a bit more intuitive, has to do with the need to handle the most common type of missing data for event histories: right-censored event times. In the next section, I describe in greater detail the different types of censoring and return to a discussion of how hazard probabilities can be easily computed, even in the presence to right-censored observations.

Censoring

Missing data is endemic to longitudinal study settings; survival analysis is no exception. The various mechanisms for missing data in the survival context are usually grouped under the encompassing term, *censoring*. Most generally, censoring occurs when the exact survival time is only known for a portion of the sample, with event times for the remaining subjects only known to occur in certain intervals. There are three categories of censoring: left, interval, and right censoring. Left censoring occurs when a subject in the sample has experienced the event of interest prior to the onset of observation. For example, in the NLSY97 data, the first wave of data collection occurred in 1997. Respondents who had their first drink before that first wave would be left-censored in 1997. Interval censoring occurs when a subject is only known to have experienced the event of interest within a given time interval but the exact time is unknown. For example, in the NLSY97 study, data collection occurred every year; however, some respondents would not be available at every follow-up. A respondent who had not had his first drink in 1997, was missing for the 1998–2000 follow-ups, and then reported in 2001 having had his first drink of alcohol since the date of his last interview, would be interval-censored between 1997 and 2001. Right censoring occurs when a subject in the sample has not experienced the event of interest at the cessation of observation. In this case, all

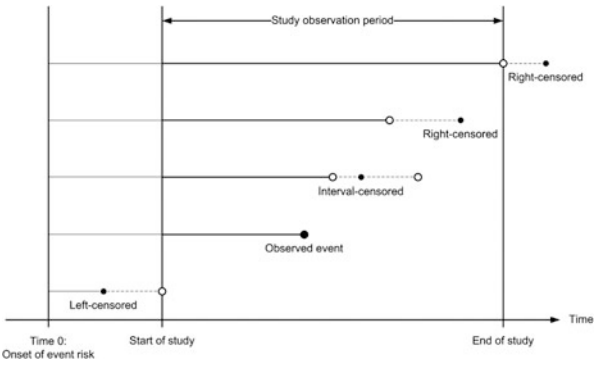

Fig. 22.1 Timeline diagram of complete and censored event histories with *filled circles* representing actual event times and *open circle* representing censoring times

that is known about the event timing is that it occurs sometime after the last observation point. For example, if a subject had not had her first drink by 2002 and then was lost to follow-up for all remaining waves, she would be right-censored at 2002. Figure 22.1 provides a timeline graphic illustrating the three different types of censoring.

As previously noted, censoring is simply the general term used for missing data mechanisms in survival data. The range of assumptions about the censoring parallel the more general missing data assumptions applied to other data settings. The assumption of noninformative censoring corresponds to the assumption of ignorable missingness. If the distribution of censoring times is independent of event times, conditional on the set of observed covariates, then the censoring may be treated as noninformative. The case of informative censoring corresponds to nonignorable missingness. In these situations, censoring times depend upon event times. For this chapter, I focus on the most common censoring in prospective longitudinal studies: random right censoring.

The Record of Event Histories

The "What," "When," and "Who"

In order to describe and model an event history or survival duration process, one must determine the answers to three basic questions: What? Who? When? *What* is the event, i.e., what constitutes an event occurrence? *When* is someone initially at risk for the event? *Who* is at risk for the event? Although these questions may seem trivial or self-evident, they must be made explicit as they have direct bearing on how the event history data is constructed and subsequently analyzed.

There are different types of survival events. The simplest is a single, nonrecurring or non-repeatable event. This means the event of interest is a singular event that can occur once and only once for each person and that after an individual has experienced the event, there is no further "risk" to the individual. Given the

historical development of survival models in the area of life table analysis, it is not surprising that the main focus for methods development has been around single, terminating events. This chapter focuses on techniques for this classic event type. For the data example, onset or initiation of alcohol use is a single, nonrecurring event in that any given individual can only have her first drink one time in her life.

There can be time-to-event processes, particularly in fields of social science research, which do not fit the single event model. Most generally, such data can be referred to as *multivariate* survival or event history data. In the survival analysis literature, multivariate event processes are typically categorized as either multiple event histories or recurrent event histories. Analysis of type of multivariate survival data, multiple event or multiple process, is beyond the scope of this chapter but see Hougaard (2000) for more information. However, elements of multivariate process can reduce to single, nonrecurring events, e.g., time-to-arrest is recurring but time-to-first-arrest is singular.

In addition to characterizing the type of event, it is also important to clarify what constitutes an event occurrence. With the traditional event of death, the determination of event occurrence is indubitable. For some events, especially when ascertainment of event occurrence requires self-report, a consistent definition of the event must be applied. For the data example, alcohol use onset is characterized by having one's first drink, where "drink" is defined by the NLSY97 survey prompt as "a can or bottle of beer, a glass of wine, a mixed drink, or a shot of liquor" and respondents are instructed, "do not include childhood sips that you might have had from an older person's drink." There are also events that may not be directly observable. For example, for an event such as the onset of a major depressive episode, a measurement model or set of diagnostic criteria must be utilized to determine event occurrence. For more on survival models that incorporate uncertainty or measurement error in event occurrence, see Masyn (2008).

In order to measure time-to-event, there must be a time zero for each individual—the time at which each individual is initially at risk for the event. For many developmental studies, birth can reasonably serve as the time zero for life course events. In these situations, the time scale is chronological age. Risk begins at birth and the timing of the event is recorded as the age of the individual at the time the event occurs. For some event processes, a different time scale is more appropriate. For example, time-to-divorce is best measured using date of marriage as time zero. Since it is rare that the beginning of a study corresponds to an appropriate time zero and since it is also unusual for the beginning of the study to be the same distance, time-wise, from each participant's own time zero, most event times recorded in calendar years or in terms of study timeline must be centered according to the relationship between each participant's time zero and the calendar time of the event. For the NLSY97 example, time zero is defined as birth since it is the *age* of onset of alcohol use that is to be modeling. To convert event records from calendar time to age chronology, it is first necessary to determine each respondent's age at the time of the survey interviews. Once each participant's age at each interview date is computed, then any onsets of alcohol use occurring, say, between the 1997 and 1998 interviews, can be translated to onsets occurring in a specific age

range. For example if a participant is 12 years old at the time of the 1997 interview and 13 years old at the 1998 interview and initiates alcohol use between the Wave 1 and Wave 2 interviews, the onset age for that individual is between age 12 and 13.

Deciding on the "who" is similar to specifying the target population for any quantitative study—the population to which the researcher would ideally like to generalize the results. The key characteristic for the target population for an event history analysis is that everyone in the target population must be, at some point in time, at risk for the event of interest. The initial *risk set* in the observed sample consists of all individuals at risk for the event at a designated time zero. In a discrete-time event history analysis, the risk set changes for each time interval. The initial risk set should be all cases in the sample representative of the target population. For a single, nonrecurring event, such as onset of alcohol use, once an individual experiences the event of interest, she is not in the observed risk set for any subsequent time intervals. If an individual is lost to follow-up prior to experiencing the event (i.e., is right-censored), she is not in the observed risk set for any subsequent time intervals. Only those individuals in the risk set for a given time interval contribute information to the estimation of the hazard function for that time interval.

Constructing Discrete-Time Event Histories from Prospective Longitudinal Panel Data

Most survival analysis textbooks begin with random, right-censored observed event histories in the following format: (Time_i,Event_i), where Time_i is the last time period for which individual i was under observation, and Event_i is an event indicator such that $\text{Event}_i = 1$ if individual i experienced the event during time period Time_i (i.e., $T_i = \text{Time}_i$) and $\text{Event}_i = 0$ if individual i was right-censored during the time period immediately following Time_i. For drinking onset using 1-year age intervals, an individual who onset at age 15 would have ($\text{Time}_i = t_{15}$, $\text{Event}_i = 1$), while an individual who did not have a first drink by age 15 but was lost to follow-up before age 16 would have ($\text{Time}_i = t_{15}$, $\text{Event}_i = 0$). One of the challenges when using survival analysis techniques in a real data setting is that the raw data from prospective longitudinal panel studies do not usually appear in this (Time,Event) form. It is usually necessary to construct event histories for each individual in the data set, as is the case with alcohol onset survival data in the NLSY97. At the first wave, all individuals were asked if they had *ever* had a drink of alcohol. Those answering "yes" were left-censored at the beginning of the study. Those answering "no" skipped the other survey questions related to alcohol use. At Wave 2 (1998), those same respondents (who had not had a first drink by Wave 1) were asked whether or not they had a drink of alcohol since the date of the last interview. If they answered "yes," then they had their first drink in the time interval between the Wave 1 and Wave 2 interviews. Constructing event histories for the onset age of

alcohol use from data involves tracking each individual's responses to the alcohol-related survey questions across all of the available waves and integrating the event time information contained in those responses with the chronological age of the respondents at each wave. Table 22.1 illustrates, for 11 of the many possible response patterns across the waves, the translation from survey responses to event histories.

Once the event histories for each individual have been constructed from the raw panel data, event history indicators must be created that will serve as the actual dependent variables in the hazard regression model. There is a binary event history indicator for each of the J time periods under observation, $(e_{1i}, e_{2i}, \ldots, e_{Ji})$. The event history indicator corresponding to the time period during which an individual is known to have experienced the event is coded as 1; event history indicators corresponding to all prior time periods (before the event) are coded as 0. The event history indicators for time periods following the event are coded as missing as are all indicators following a right censoring. In general, the relationship between the $(\text{Time}_i, \text{Event}_i)$ data and the event indicators

$$e_{ji} = \begin{cases} 1 & \text{if} \quad \text{Time}_i = t_j, \quad \text{Event}_i = 1, \\ 0 & \text{if} \quad \text{Time}_i > t_j \quad \text{or} \quad \text{Time}_i = t_j, \quad \text{Event}_i = 0, \\ . & \text{if} \quad \text{Time}_i < t_j. \end{cases} \tag{22.5}$$

For the NLSY97 example, there are a total of 11 time periods; the first time period is birth to age 12.5 years (the approximate mean age for the age 12 interview) and the remaining 10 time periods are 1 year in length. For an individual i who has his first drink between the age 15 and 16 interviews, ($\text{Time}_i = t_5 = [15.5, 16.5)$, $\text{Event}_i = 1$). The event indicator values are all 0 for e_{1i} through e_{4i}, 1 for e_{5i}, and missing for e_{6i} through e_{11i}. For an individual i who has not had his first drink by the age 16 interview and is missing (e.g., lost to follow-up) for the age 17 interview, ($\text{Time}_i = t_5 = [15.5, 16.5)$, $\text{Event}_i = 0$). The event indicator values are all 0 for e_{1i} through e_{5i} and missing for e_{6i} through e_{11i}. Table 22.2 presents all possible indicator patterns for observed and right-censored onset ages for the NLSY97 example along with the corresponding frequencies and percentages of cases in the data with matching event histories.

Modeling Event Histories

Estimating the Unconditional Hazard Function

The most straightforward method for estimation of the unconditional survival probabilities for complete or right-censored data is to first estimate the hazard probabilities. Essentially, the hazard probability is estimated for each time interval by taking the ratio of the number of subjects experiencing the event in a given

Table 22.1 Example NLSY97 survey responses for the first seven waves (1997–2003) with corresponding event histories for onset age of alcohol use

Response pattern	1	2	3	4	5	6	7	8	9	10	11
1997: Ever had a drink?	Yes	No	No	No	No	No	No	No	No	No	No
1998: Had a drink since date of last interview?		Yes	No	No	No	No	No	No	No	No	No
1999: Had a drink since date of last interview?			Yes	No	No	No	No	No	No	No	No
2000: Had a drink since date of last interview?				Yes	No	No	No	No	Missing	Missing	No
2001: Had a drink since date of last interview?					Yes	No	No	No	Missing	Missing	Missing
2002: Had a drink since date of last interview?						Yes	No	No	No	Yes	Missing
2003: Had a drink since date of last interview?							Yes	No	Yes		Missing
Event time (in calendar years)	Before 1997 interview	Between 1997 and 1998 interviews	Between 1998 and 1999 interviews	Between 1999 and 2000 interviews	Between 2000 and 2001 interviews	Between 2001 and 2002 interviews	Between 2002 and 2003 interviews	After 2003 interview	Between 2002 and 2003 interviews	Between 1999 and 2002 interviews	After 2000 interview
Censored?	Left	No	No	No	No	No	No	Right	No	Interval	Right
Onset age[a]	0–Age_{97}	Age_{97}–Age_{98}	Age_{98}–Age_{99}	Age_{99}–Age_{00}	Age_{00}–Age_{01}	Age_{01}–Age_{02}	Age_{02}–Age_{03}	>Age_{03}	Age_{02}–Age_{03}	Age_{99}–Age_{02}[b]	>Age_{00}

[a]Age_{xx} is the age of the respondent (in years) at the time of the 19*xx* (or 20*xx*) interview

[b]For the analysis in this chapter, interval-censored observations such as this were treated as right-censored at the beginning of the interval, e.g., onset age > Age_{99}

Table 22.2 Data coding and frequency descriptives for NLSY97 event history indicators ($n = 8{,}280$)

	j	t_j	Time$_i$	Event$_i$	e_{1i}	e_{2i}	e_{3i}	e_{4i}	e_{5i}	e_{6i}	e_{7i}	e_{8i}	e_{9i}	e_{10i}	e_{11i}	f	%
Age[a] at first drink	1	[0,12.5)	t_1	1	1											1,390	16.8
	2	[12.5,13.5)	t_2	1	0	1										724	8.7
	3	[13.5,14.5)	t_3	1	0	0	1									848	10.2
	4	[14.5,15.5)	t_4	1	0	0	0	1								844	10.2
	5	[15.5,16.5)	t_5	1	0	0	0	0	1							576	7.0
	6	[16.5,17.5)	t_6	1	0	0	0	0	0	1						489	5.9
	7	[17.5,18.5)	t_7	1	0	0	0	0	0	0	1					420	5.1
	8	[18.5,19.5)	t_8	1	0	0	0	0	0	0	0	1				305	3.7
	9	[19.5,20.5)	t_9	1	0	0	0	0	0	0	0	0	1			177	2.1
	10	[20.5,21.5)	t_{10}	1	0	0	0	0	0	0	0	0	0	1		149	1.8
	11	[21.5,22.5)	t_{11}	1	0	0	0	0	0	0	0	0	0	0	1	46	0.6
Pre-onset last age[a] observed (prior to right censoring)	1	[0,12.5)	t_1	0	0											203	2.5
	2	[12.5,13.5)	t_2	0	0	0										264	3.2
	3	[13.5,14.5)	t_3	0	0	0	0									278	3.4
	4	[14.5,15.5)	t_4	0	0	0	0	0								253	3.1
	5	[15.5,16.5)	t_5	0	0	0	0	0	0							227	2.7
	6	[16.5,17.5)	t_6	0	0	0	0	0	0	0						124	1.5
	7	[17.5,18.5)	t_7	0	0	0	0	0	0	0	0					111	1.3
	8	[18.5,19.5)	t_8	0	0	0	0	0	0	0	0	0				272	3.3
	9	[19.5,20.5)	t_9	0	0	0	0	0	0	0	0	0	0			228	2.8
	10	[20.5,21.5)	t_{10}	0	0	0	0	0	0	0	0	0	0	0		150	1.8
	11	[21.5,22.5)	t_{11}	0	0	0	0	0	0	0	0	0	0	0	0	202	2.4
$\hat{P}_h(t_j) = \hat{\Pr}(e_j = 1)$					0.015	0.108	0.149	0.185	0.166	0.183	0.204	0.199	0.186	0.272	0.185		

[a]"Age" represents the time interval between consecutive study wave interviews, e.g., *age at first drink* = [13.5,14.5) indicates onset occurred between the age 13 and 14 interviews; *pre-onset last age observed* = [14.5,15.5) indicates loss to follow-up occurred after the age 15 interview

Table 22.3 Event history frequency statistics with sample hazard and survival probabilities

j	t_j = Age[a] (in years) at first drink	# at risk	# events	$\hat{P}_{\mathrm{h}}(t_j)$	$\hat{P}_{\mathrm{S}}(t_j)$
1	[0,12.5)	8,280	1,390	0.015	0.832
2	[12.5,13.5)	6,687	724	0.108	0.742
3	[13.5,14.5)	5,699	848	0.149	0.632
4	[14.5,15.5)	4,573	844	0.185	0.515
5	[15.5,16.5)	3,476	576	0.166	0.430
6	[16.5,17.5)	2,673	489	0.183	0.351
7	[17.5,18.5)	2,060	420	0.204	0.280
8	[18.5,19.5)	1,529	305	0.199	0.224
9	[19.5,20.5)	952	177	0.186	0.182
10	[20.5,21.5)	547	149	0.272	0.133
11	[21.5,22.5)	248	46	0.185	0.108

[a]"Age" before 12 are based on self-reported age of first drink; "age" values after 12 are determined by the time interval between consecutive study wave interviews, e.g., "at risk" at age = [13.5,14.5) indicates onset did not occur prior to the age 13 interview; "event" at age = [13.5,14.5) years indicates a first drink occurring between the age 13 and 14 interviews

interval over the number of subjects observed to be "at risk" for the event at the beginning of the given time interval. A subject is observed to be at risk in time period j if she has not experienced the event prior to period j and is not right-censored before the end of time period j.

Let n_j be the number observed to be at risk during time period t_j (equal to the number of non-missing e_j values in the sample) and let d_j be the number of observed events during time period (equal to the number of individuals in the sample with $e_j = 1$). The estimated hazard probability for t_j is then calculated as

$$\hat{P}_{\mathrm{h}}(t_j) = \hat{\Pr}(e_j = 1) = \frac{d_j}{n_j}. \tag{22.6}$$

Using the relationship between the hazard probability and the survival probability given in (22.4), the estimated survival probability for t_j is given by

$$\hat{P}_{\mathrm{S}}(t_j) = \prod_{m=1}^{j} \left(1 - \hat{P}_{\mathrm{h}}(t_m)\right). \tag{22.7}$$

For the NLSY97 data illustration, the estimated hazard probabilities for onset of alcohol use and the corresponding survival probabilities are given in Table 22.3. For example, in the time period between the age 13 and 14 interviews (t_3), 5,699 respondents of the original 8,280 are still at risk for onset, since 1,390 + 724 = 2,114 had experienced the event prior to t_3 and 467 were right-censored prior to t_3, leaving 8,280 − (2,114 + 467 = 5,699) in the t_3 risk set (i.e., 5,699 individuals in the sample have non-missing e_3 values). During that time period, 848 initiated alcohol use (i.e., 848 individuals in the sample have $e_3 = 1$), yielding an estimated hazard probability of

$\hat{P}_{\mathrm{h}}(t_3) = \hat{\Pr}(e_3 = 1) = 848/5,699 = 0.149$. Note that the estimated unconditional probabilities for the event indicators given on the last row of Table 22.2 match the estimated hazard probabilities given in Table 22.3.

The probability of survival beyond that time period is then the product of the complements of the hazard probabilities for the first three time periods:

$$\begin{aligned}\hat{P}_{\mathrm{S}}(t_3) &= \left(1 - \hat{P}_{\mathrm{h}}(t_1)\right) \cdot \left(1 - \hat{P}_{\mathrm{h}}(t_2)\right) \cdot \left(1 - \hat{P}_{\mathrm{h}}(t_3)\right) \\ &= (1 - 0.015) \cdot (1 - 0.108) \cdot (1 - 0.149) = 0.632.\end{aligned}$$

Figure 22.2 displays the plots for the estimated hazard and survival probabilities for the data example. The hazard probabilities are plotted at the midpoint of each time interval using an approach similar to that used to construct histograms—that the area for a bar spanning a given time period and not the height of the bar reflects the value of the hazard probability for that time period. The hazard probabilities increase over adolescence with the highest hazards between the age 17 and 18 interviews, with approximately 20 % of those still at risk prior to that time period initiating alcohol use, and between the age 20 and 21 interviews, with approximately 27 % of those still at risk prior to that time period having their first drink between the age 20 and 21 interviews. By the age 22 interview, only 11 % of respondents have not yet had their first drink of alcohol (i.e., "survived" to the age 22 interview).

In addition to estimated hazard probabilities and the corresponding survival probabilities, another useful summary statistic sometimes used for event history analysis is the median survival time. This is the time period at which points approximately 50 % "survived" the event. That is,

$$S_{0.50} = t_j : P_{\mathrm{s}}\left(t_{j-1}\right) > 0.50 \qquad \text{and} \qquad P_{\mathrm{s}}\left(t_j\right) \leq 0.50. \tag{22.8}$$

The dotted lines drawn in Fig. 22.2b locate the median time which can also be obtained from the information in Table 22.3: $S_{0.50} = t_5 = [15.5, 16.5)$.

Estimating the Conditional Hazard Function

Although describing the unconditional hazard and survival functions for a given event process is useful, researchers often want to further examine the effects of risk and protective factors as well as preventive interventions on the time-to-event. In order to relate event time to hypothesized predictors, the hazard function remains the analytic quantity of interest as it allows the inclusion of right-censored as well as complete event histories. Since the hazard function is a set of probabilities, corresponding to the event history indicator probabilities for each time period, a natural modeling approach is one most commonly used for binary observed outcomes: logistic regression.

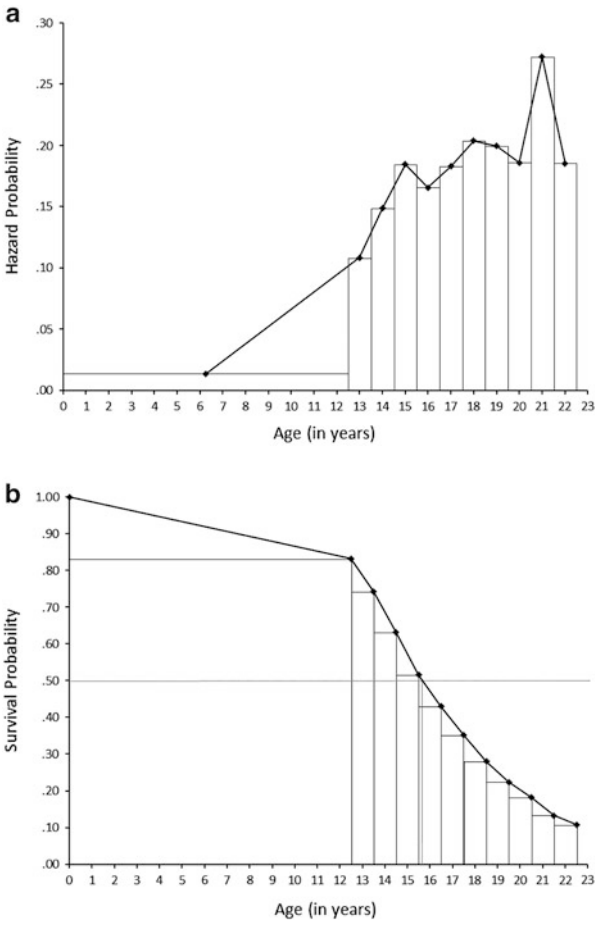

Fig. 22.2 Unconditional model-estimated (**a**) hazard probabilities and (**b**) corresponding survival probabilities for onset age (in years) of alcohol use. (*Note*: Estimated hazard probabilities in (**a**) corresponding to the first time interval [0,12.5) is plotted so that the area under the bar is equal to the hazard probability.)

In an unconditional model without any predictors, the logit hazard regression would be written as

$$\text{logit}\left(P_{\text{h}}\left(t_j\right)\right) = \log\left(\frac{P_{\text{h}}\left(t_j\right)}{1 - P_{\text{h}}\left(t_j\right)}\right) = \log\left(\frac{\Pr\left(e_j = 1\right)}{1 - \Pr\left(e_j = 1\right)}\right) = \alpha_j, \qquad (22.9)$$

where α_j is the log hazard odds of the event[1]. The inverse logit of the estimate for α_j is identical to the sample hazard probabilities as given in (22.6); that is,

[1] When identifying the event history indicators as "categorical" in the VARIABLE command, the Mplus software estimates a "threshold" parameter, τ_j, that is the additive inverse of α_j; that is, $\hat{\alpha}_j = -\hat{\tau}_j$.

$$\hat{P}_{\mathrm{h}}(t_j) = \hat{\Pr}(e_j = 1) = \frac{d_j}{n_j} = \frac{1}{1 + \exp(-\hat{\alpha}_j)}. \tag{22.10}$$

Extending (22.9) to include observed predictors of the time-to-event, the logit hazard is specified as a linear function of a time-invariant covariate(s), x, and time-varying covariate(s), w_j, as given by

$$\begin{aligned} \operatorname{logit}\left(P_{\mathrm{h}}(t_j)\right) &= \log\left(\frac{P_{\mathrm{h}}(t_j)}{1 - P_{\mathrm{h}}(t_j)}\right) = \log\left(\frac{\Pr(e_j = 1)}{1 - \Pr(e_j = 1)}\right) \\ &= \alpha_j + \beta_j x + \gamma_j w_j, \end{aligned} \tag{22.11}$$

or, alternatively expressed as

$$P_{\mathrm{h}}(t_j) = \Pr(e_j = 1) = \frac{1}{1 + \exp\left(-\left(\alpha_j + \beta_j x + \gamma_j w_j\right)\right)}, \tag{22.12}$$

where β_j is the difference in the log hazard odds and $\exp(\beta_j)$ is the hazard odds ratio (hOR) of the event occurring in t_j corresponding to a positive one unit difference in x, controlling for w_j, and $\exp(\gamma_j)$ is the hOR of the event occurring in t_j corresponding to a positive one unit difference in w_j, controlling for x. A positive log hOR (e.g., $\beta_j > 0$) equates to a hOR greater than one (e.g., $\exp(\beta_j) > 1$) meaning that higher values of the predictor correspond to *higher* hazard odds which translates to *shorter* median survival times. For example, if a family history of alcohol abuse was positively associated with the logit hazard of drinking onset, then individuals with a family history of alcohol abuse would have earlier onset ages, on average.

Notice that in the specification given in (22.11), both the time-invariant covariate, x, and the time-varying covariate, w, have time-varying effects, β_j and w_j, respectively. One common model nested within the model in (22.11) constrains all covariate effects to be time-invariant, that is,

$$\operatorname{logit}\left(P_{\mathrm{h}}(t_j)\right) = \alpha_j + \beta x + \gamma w_j. \tag{22.13}$$

The difference in the logit hazard values corresponding to a one unit differences in x does not depend on j, meaning that the effect of x on the logit hazard is *constant* over time. Since the hOR for x is constant over time, the hazard odds are proportional over time (i.e., have the same ratio); thus, this specification is known as the proportional hazard odds model.

Types of covariates. Covariates can be either continuous or categorical. Categorical predictors enter the model via a set of dummy variables or other set of contrast variables just as they would in a regular linear regression. The logit hazard model assumes that the log hazard odds change linearly across the range of a given continuous covariate but that linearity assumption can be relaxed, e.g., including

polynomial functions of the covariates such as adding x^2 as an additional predictor. To evaluate the moderation of one covariate effect by another, an interaction term between the two covariates is included as an additional predictor in the model.

Covariate predictors of the event history process can be time-varying or time-variant. Each covariate can be specified to have time-varying or time-invariant effects. Time-invariant covariates are covariates whose values remain constant for each individual during the full period of observation. Such explanatory variables might include individual attributes, such as race, or variables whose values are fixed at or before time zero, such as treatment status. Time-varying or time-dependent covariates have values that change over time for the individual but are assumed to be constant within a given discrete-time interval. Although the model specification in (22.11) makes the inclusion of time-dependent covariates a straightforward matter, the implications for the interpretation of the model are much more complex. As with other longitudinal models, inclusion of time-dependent covariates introduces the issue of reciprocal causation (Cox & Oakes, 1984; Singer & Willett, 2003). Essentially, when linking contemporaneous information on a predictor and outcome, the directional arrow, that is, which variable is doing the influencing, is impossible to infer from the data for certain types of covariates whose values could be influenced by the co-occurring outcome. One method for side-stepping this issue is to time-lag the covariate values. That is, to have the covariate value in a previous time period, e.g., $j - 1$, or the change in the covariate value across two preceding time periods, e.g., $j - 2$ to $j - 1$, predicting event occurrence in time period j.

Maximum likelihood estimation. The estimate for the unconditional hazard probability given in (22.6) is equal to the maximum likelihood estimate for complete and noninformative right-censored discrete-time data. The model given in (22.11) can be estimated using full-information maximum likelihood (FIML) estimation under the MAR assumption (corresponding to the assumption of noninformative right censoring) (Muthén & Masyn, 2005; Singer & Willett, 2003).

Model testing. There are two more standard asymptotic statistical inference tests for evaluating the statistical significance of single covariates in a maximum likelihood estimated logit hazard regression. One is the Wald test and the other is the likelihood ratio test (LRT).

Although the Wald and LRT are asymptotically equivalent, the LRT tends to be preferred over the Wald test.

The LRT can be extended to multivariate case to test multiple parameters at once. In fact, the LRT can be used to evaluate the difference in fit between any two nested models. Most generally the likelihood ratio test statistic (LRTS) is computed as follows:

$$\text{LRTS} = -2[\text{LL}_0 - \text{LL}_1], \tag{22.14}$$

where LL_0 is the maximum log likelihood for the estimated null model nested within the alternative model with a maximum log likelihood of LL_1. The LRTS is approximately chi-squared distributed under the null model with $\text{df} = \text{npar}_1 - \text{npar}_0$, where npar_1 and npar_0 are the number of freely estimated parameters in the

alternative and null models, respectively. A significant p-value leads to rejection of the null model in favor of the alternative; i.e., a significant p-value suggests that the alternative model fits the data statistically significantly better than the more restrictive null model. The LRT can be used for inferences about many different aspects of the hazard model including the significance of a single covariate (continuous or categorical) and the time-dependency of the effects of a given covariate.

Statistical power. With regards to issues of power and sample size, one may draw from the field of categorical data analysis. In the discrete-time survival setting, the necessary sample size to fit specific models or the power to detect certain covariate effects is related to the rarity of the event relative to the sample size. Peduzzi, Concato, Kemper, Holford, and Feinstein (1996) demonstrate the "rule of 10," showing that at least ten events per parameter per time interval are necessary to obtain reliable estimate of logistic hazard regression coefficients and their standard errors.

Data Example Results

For all conditional models, the effects of the covariates on the hazard odds for the first time period [0, 12.5) were allowed to be different than the effects on the hazard odds for the other ten time periods, each 1 year in length. Beginning with the main effects model, allowing all covariates to have time-varying effects across all 11 time periods, the addition of the covariates makes a statistically significant improvement over the model with no covariates (adj. LRTS = 259.70, df = 66, $p < 0.001$)[2]. Testing for interaction effects across the time periods, only the addition of two-way interactions between the parent education covariate and the remaining covariates resulted in a statistically significant improvement in model fit (adj. LRTS = 42.64, df = 10, $p < 0.001$). The model results for this final model are presented in Table 22.4. For the hazard odds of initiating alcohol use prior to the age 12 interview, females have lower odds compared to males at mean levels of parent education, controlling for race and urban residency (Est. hOR = 0.72, $p < 0.001$). Equivalently stated, males have approximately 1/0.72 = 1.39 times higher hazard odds for initiating alcohol use prior to age 12.5.[3] This effect stands in contrast to the nonsignificant difference in the hazard odds for females compared to males for the remaining ten time periods, from ages 12.5 to 22.5. There is a significant but small interaction between parent education and gender such that the hazard odds for females is slightly higher than males among respondents with

[2] The adjusted likelihood ratio test statistic is computed using the log likelihood values and scaling correction factors obtained with the "MLR" estimator in Mplus, the default maximum likelihood estimator for binary and ordinal outcomes.

[3] Recall this is self-reported age of onset prior to the age 12 interview. It may be that males are more likely to report having already initiated alcohol use at the baseline interview regardless of their actual onset status.

Table 22.4 Model results for final SEM event history model (LL = −16,505.55, number of parameters = 33, n = 8,280)

e_1 on	Est.	SE	p	e_2, …, e_{11} on	Est.	SE	p
Female	**−0.33**	**0.08**	**<0.01**	Female	0.04	0.04	0.36
Urban	0.18	0.10	0.06	**Urban**	**0.15**	**0.05**	**0.01**
Black	**−0.46**	**0.10**	**<0.01**	**Black**	**−0.71**	**0.06**	**<0.01**
Hispanic	**−0.29**	**0.12**	**0.01**	**Hispanic**	**−0.27**	**0.07**	**<0.01**
Other	−0.03	0.23	0.91	**Other**	**−0.78**	**0.13**	**<0.01**
Parent Education[a]	**−0.08**	**0.04**	**0.03**	**Parent Education**[a]	**−0.05**	**0.02**	**0.01**
Female × ParEd	0.04	0.03	0.14	**Female × ParEd**	**0.03**	**0.01**	**0.05**
Urban × ParEd	−0.01	0.04	0.71	Urban × ParEd	0.01	0.02	0.61
Black × ParEd	**0.10**	**0.04**	**0.02**	**Black × ParEd**	**0.09**	**0.02**	**<0.01**
Hispanic × ParEd	**0.12**	**0.03**	**<0.01**	**Hispanic × ParEd**	**0.06**	**0.02**	**<0.01**
Other × ParEd	−0.01	0.08	0.91	**Other × ParEd**	**0.11**	**0.03**	**<0.01**
Intercept	*Est.*	*SE*	p	Intercept	*Est.*	*SE*	p
e_1	1.33	0.09	<0.01				
e_2	1.98	0.06	<0.01	e_7	1.06	0.08	<0.01
e_3	1.59	0.06	<0.01	e_8	1.08	0.09	<0.01
e_4	1.26	0.06	<0.01	e_9	1.22	0.11	<0.01
e_5	1.42	0.07	<0.01	e_{10}	0.55	0.13	<0.01
e_6	1.17	0.07	<0.01	e_{11}	1.02	0.22	<0.01

[a]Parent Education (ParEd; average of highest grade ever completed by each resident parent) is centered at 12th grade

parent education values above the mean and is slightly lower than males among respondents with parent education values one grade or more below the mean. There is no evidence of an interaction between urban residency and parent education for either the first time period or the last ten time periods; however, there is some evidence of a slightly increased hazard odds for respondents living in urban areas compared to those living in nonurban area during ages 12.5–22.5 (Est. hOR = 1.16, p = 0.01). The largest and most interesting effects are the interactions between race and parent education. Figure 22.3 depicts the estimated hazard probabilities and corresponding survival probabilities for the four race groups (White, Black or African American, Hispanic, and other) at the 10th, 50th, and 90th percentiles of parent education (9th grade, 12th grade, and 16th grade). At the lower levels of average parent education, the hazard probabilities for White respondents are much higher than any of the other groups, with the Black and "other" group being comparably low. This translates to the lowest median survival time for Whites, followed by Hispanics, as indicated by the dotted lines in Fig. 22.3b. For increasing values of parent education, the hazard probabilities for White and Hispanic students decrease slightly and become much more similar to each other but still higher than the Black and "other" group. In other words, there are marked differences between the four race groups with respect to the hazard odds of initiating alcohol use between the ages 12.5 and 22.5 among those with lower average parent education levels but those racial differences in drinking onset hazards narrow at higher levels

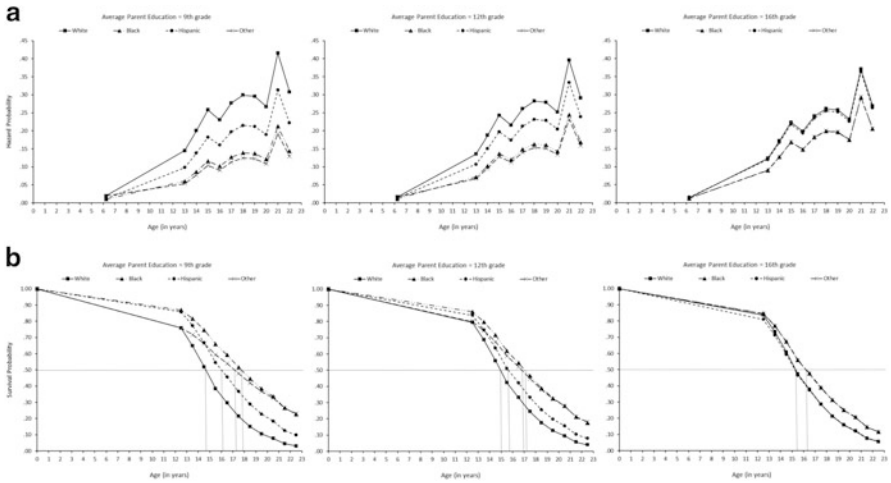

Fig. 22.3 Conditional model-estimated (**a**) hazard probabilities and (**b**) corresponding survival probabilities for onset age (in years) of alcohol use by average residential parent education (higher grade completed) and race/ethnicity, controlling for gender and urban area residence

of mean parent education, partly due to the fact that the hazard odds of drinking decreases among White respondents across increasing levels of mean parent education (Est. hOR = 0.95, $p = 0.01$). The same narrowing can be seen in the decreasing distances between the median survival times shown by the dotted lines on the plots in Fig. 22.3b.

Conclusion

This chapter provided an introduction to the general set of statistical methods developed to model discrete-time event histories. The approach for constructing and modeling single event discrete-time data was detailed and illustrated using data on the onset age of alcohol use in adolescence and early adulthood from the first seven waves of NLSY97. The models herein can be done in any statistical software that performs logistic regression. However, placing the discrete-time event history model in broader latent variable modeling framework allows for an impressive number of model extensions that are not possible within the bounds of conventional logistic regression. In what follows, I describe how the event history indicators can be used as a set of binary outcome variables in different latent variable model specifications and then highlight some of the powerful extensions enabled by these different specifications.

Discrete-Time Survival Analysis as a Structural Equation Model

In modeling repeated measures growth data, such as trajectories of aggressive-disruptive behaviors in childhood, the two dominant modeling approaches are (1) a multilevel random effects model with occasions nested within individuals and (2) a latent growth model with the repeated measures on an individual treated as multiple measures of person-level random variables specified as a structural equation model. An analogous pair of approaches exists for discrete-time survival analysis as well. One approach is to create a person-period data set, "stacking" the event history indicators, and then estimating the logit hazard regression as a hierarchical generalized linear model (HGLM) with time periods nested within individuals. The other approach is to create a person-level data set with the event history indicators in "wide" format, and then estimating the logit hazard regression as a structural equation model with J independent, binary endogenous variables, related to observed predictors using a logit link function. Figure 22.4a displays a path diagram of the conditional path analysis model with the event indicators as defined in (22.5) and the relationships between the observed exogenous variables and the event indicators given in (22.11).

Alternative model specifications. Utilizing the broader latent variable framework within which the model depicted in Fig. 22.4a falls, there are alternative model specifications for logit hazard regressions estimated with FIML that, in the case of a single, nonrecurring event history, yield equivalent models, i.e., models that have the same number of free parameters and yield identical maximum likelihood values. Figure 22.4 depicts three equivalent models: (a) a multivariate logistic regression path analysis model; (b) a structural equation model with a single latent factor for which all factor loadings for the event indicators are fixed at one and the factor variance is fixed at zero; and (c) a 1-class finite mixture model with the event indicators as the latent class measures. Figure 22.5 depicts two equivalent specifications for a hazard regression model with time-invariant effects of a time-invariant covariate: (a) a multivariate logistic regression model with the logit hazard regression coefficient for x constrained to be equal for all $j = 1, \ldots, J$, and (b) a single factor model in which the factor is regressed on x, all factor loadings for the event indicators are fixed at one, and the factor residual variance is fixed at zero.

In modeling single-occurrence events in discrete-time with observed predictors, there is no reason to choose one specification over the others although I favor the model as presented in Fig. 22.4a over the non-varying latent variables (either a factor with zero variance or a mixture with only one latent class) as the inclusion those latent variables is unnecessary. However, building on these alternative model specifications and others (including a hybrid factor mixture model) enable different extensions beyond single events including unobserved population heterogeneity (frailty) in the event history process, latent variable predictors of event history, recurring events, competing risks, parallel and sequential growth processes, onset-to-growth models, parallel event history processes, measurement error on event

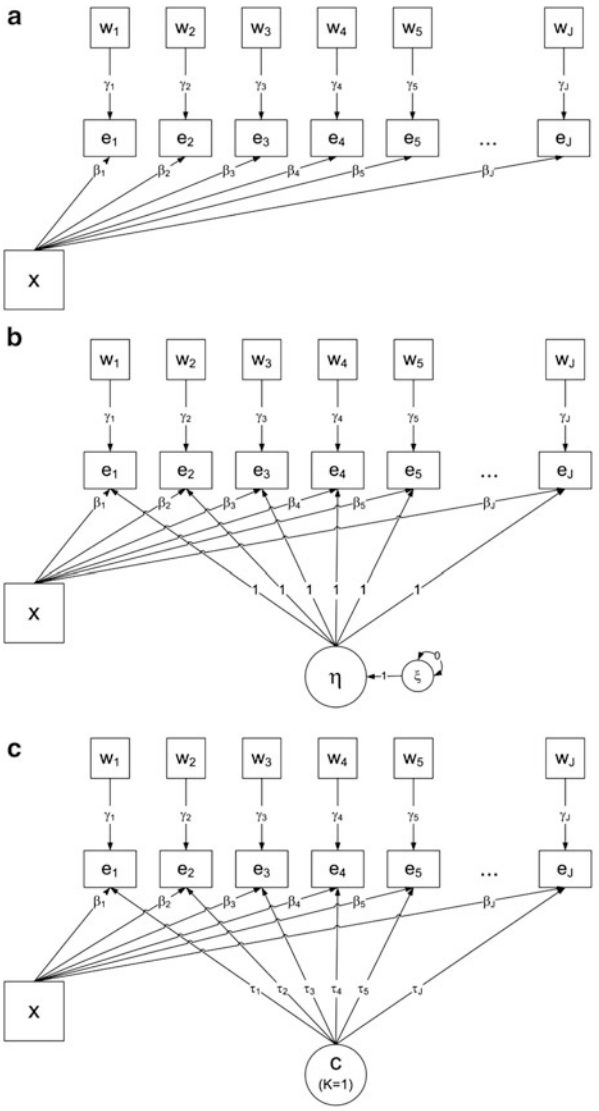

Fig. 22.4 Path diagrams for three equivalent model specifications for single, nonrecurring event history analysis with time-invariant (x) and time-varying (w_j) covariates in a latent variable framework using J binary event history indicators as endogenous variables: (**a**) multivariate logistic regression path analysis model; (**b**) structural equation model with zero-variance latent factor with unit factor loadings; and (**c**) finite mixture model with a 1-class latent categorical variable

occurrence, hazard risk set matching, etc. And in these more complex extensions, the distinctions between these specifications and the advantages of one over another in a specific setting become clear.

While the scope of the chapter was limited to homogeneous, single, nonrecurring event processes, it should provide a foundation to pursue more advanced models that are able to accommodate increasingly complex time-to-event processes that may be of interest in particular prevention science settings; e.g., discrete-time survival mixture analysis to capture individual variability in event susceptibility over time (Muthén & Masyn, 2005; Vermunt, 2002); discrete-time survival factor

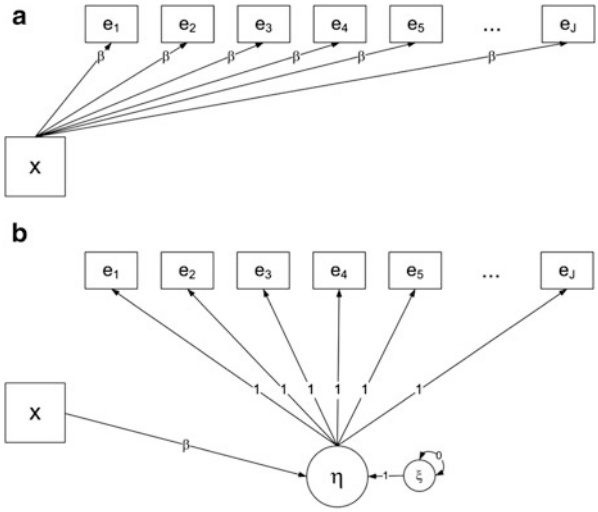

Fig. 22.5 Path diagrams for two equivalent model specifications for constraining a time-variant covariate (x) to have time-invariant effects (β) on the logit hazard function: (**a**) path model with equality constraints on the paths from x to the event history indicators; (**b**) structural equation model with zero-residual-variance latent factor with unit factor loadings regressed on x

mixture analysis modeling individual variability in recurring criminal offending (Masyn, 2009); multilevel discrete-time survival analysis modeling classroom climate influence on the hazard of individual school removal (Petras et al., 2011); multilevel discrete-time survival analysis of recurring, competing risks to model contraceptive use dynamics (Steele, Goldstein, & Browne, 2004); joint survival-growth models to examine the relationship between age of alcohol and alcohol use prevalence following onset (Malone, Northrup, Masyn, Lamis, & Lamont, 2012) or to examine the relationship between time-to-lapse and post-lapse drinking trajectories after completion of an alcohol use disorder treatment program (Witkiewitz & Masyn, 2008); joint growth-survival modeling the relationship between early aggression trajectories and the hazard of school removal (Muthén & Masyn, 2005); and parallel discrete-time survival analysis to model the reciprocal influences of the onsets of licit and illicit drug use (Malone, Lamis, Masyn, & Northrup, 2010). These examples are just a few of the ways discrete-time survival analysis can be leveraged in combination with other modeling techniques to address a range of epidemiologic, etiologic, and preventive intervention research questions of importance to prevention science.

Acknowledgement The author would like to thank Dr. Amanda Fairchild and Dr. Hanno Petras for their valuable feedback.

References

Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology, 13*, 61–98.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, 34* (2), 187–220.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data*. New York, NY: CRC Press.
Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: The Guilford Press.
Grant, B. F., & Dawson, D. A. (1997). Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. *Journal of Substance Abuse, 9*(1), 103–110.
Hingson, R. W., Heeren, T., & Winter, M. R. (2006). Age at drinking onset and alcohol dependence: Age at onset, duration, and severity. *Archives of Pediatrics and Adolescent Medicine, 160*(7), 739.
Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time to event data* (2nd ed.). Hoboken, NJ: Wiley-Interscience.
Hougaard, P. (2000). *Analysis of multivariate survival data*. New York, NY: Springer.
Malone, P., Lamis, D., Masyn, K., & Northrup, T. (2010). Modeling of the gateway drug hypothesis: A dual-process discrete-time survival analysis approach. *Multivariate Behavioral Research, 45*(5), 790–805.
Malone, P. S., Northrup, T. F., Masyn, K. E., Lamis, D. A., & Lamont, A. E. (2012). Initiation and persistence of alcohol use in United States Black, Hispanic, and White male and female youth. *Addictive Behaviors, 37*, 299–305.
Masyn, K. (2008). Modeling measurement error in event occurrence for single, non-recurring events in discrete-time survival analysis. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 105–145). Charlotte, NC: Information Age Publishing.
Masyn, K. (2009). Discrete-time survival factor mixture analysis for low-frequency recurrent event histories. *Research in Human Development, 6*, 165–194.
Morean, M. E., Corbin, W. R., & Fromme, K. (2012). Age of first use and delay to first intoxication in relation to trajectories of heavy drinking and alcohol-related problems during emerging adulthood. *Alcoholism, Clinical and Experimental Research, 36*(11), 1991–1999.
Muthén, B., & Masyn, K. (2005). Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics, 30*(1), 27–58.
Muthén, B., & Muthén, L. K. (1998–2012). *Mplus (Version 7) [Computer software]*. Los Angeles, CA: Muthén & Muthén.
Ohio State University. Center for Human Resource Research. (1997). National Longitudinal Survey of Youth, 1997. ICPSR03959-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-06-04. doi:10.3886/ICPSR03959.v2
Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*(12), 1373–1379.
Petras, H., Masyn, K., Buckley, J., Ialongo, N., Clark, M., & Kellam, S. (2011). Who is most at-risk for school removal? An application of multilevel discrete-time survival analysis to understand individual and contextual-level influences. *Journal of Educational Psychology, 103*(1), 223–237.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
Steele, F., Goldstein, H., & Browne, W. (2004). A general multilevel multistate competing risks model for event history data, with an application to a study of contraceptive use dynamics. *Statistical Modelling, 4*, 145–159.
Vermunt, J. (2002). A general latent class approach to unobserved heterogeneity in the analysis of event history data. In J. Hagenaars & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 383–407). Cambridge: Cambridge University Press.
Witkiewitz, K., & Masyn, K. E. (2008). Drinking trajectories following an initial lapse. *Psychology of Addictive Behaviors, 22*(2), 157–167.

Chapter 23
Using Mediation and Moderation Analyses to Enhance Prevention Research

Amanda J. Fairchild and David P. MacKinnon

Introduction

This chapter illustrates how mediation and moderation analyses enhance prevention and intervention work. Mediator variables describe how or why two variables are related, illustrating the means by which, or the *mechanism* by which, the variables relate. In this way, mediation effects (also called indirect effects) describe the pathway through which a predictor and an outcome relate. Early resources described mediation analysis as the "generative mechanism" through which a predictor impacts an outcome (Baron & Kenny, 1986, p. 1173). Moderator variables describe the conditions under which, or for whom, two variables relate. Thus, moderators illustrate that the relation between two variables depends on the value of a third variable. Early references described moderators as specification variables that improved the prediction of an outcome from a predictor at certain values (e.g., Saunders, 1956).

Investigations of mediating and moderating variables have the potential to guide theory-driven program evaluation in prevention research and can lend insight into the construct validity of a program. Studying mechanisms of change by investigating mediator variables directs and refines the development of evidence-based interventions because it sheds light on how an intervention achieves its effects (or alternatively why it fails to achieve effects). Mediation analysis also contributes to assessing the relative value of different behavioral determinants in multicomponent treatment packages by examining the effectiveness of individual components in multiple mediator models. Such examinations help identify the active ingredients of a program. Investigating possible moderating agents on

A.J. Fairchild (✉)
Department of Psychology, University of South Carolina, Barnwell College, Columbia, SC 29208, USA
e-mail: amanda.fairchild@sc.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_23,

intervention efficacy by evaluating moderator variables helps characterize the generalizability, or the external validity, of an intervention. Evaluation of moderator variables can also preemptively direct practice by targeting individuals who are likely to respond to treatment or by identifying individuals who may need additional services or support. Finally, analyzing mediator and moderator variables simultaneously enables other critical prevention research questions such as, "Is the process by which a program achieves its effects the same across different types of participants or in different contexts?" or "Can a mediation effect explain an interaction effect in my program?" (Fairchild & MacKinnon, 2009).

The focus of this chapter is to discuss how mediation and moderation analyses help inform program theory, development, and evaluation. To that end, we first describe and differentiate the statistical mediation and moderation models and give examples of their application. We then discuss how these models are relevant for designing and evaluating prevention programs with the ultimate goals of identifying the active ingredients of these programs and to address the question what works for whom under what conditions. The points and methods described here extend to a variety of prevention-related domains including, but not limited to, epidemiology, criminology, health promotion, child welfare, and substance use.

The Mediation Model

What Is a Mediator?

A mediator is a third variable that explains how or why two other variables are related. The mediation model posits that an independent variable, X, predicts a mediator variable, M, which in turn predicts an outcome, Y (Fig. 23.1). Thus, the mediator variable transmits the relation between an independent variable and an outcome, illustrating the mechanism through which the two variables are related. In this way, mediation analysis lends itself well to program evaluation where "causal assumptions linking program resources, activities, intermediate outcomes and ultimate goals" are considered (Wholey, 1987, p. 78).

The application of mediation is particularly relevant in prevention research, where researchers aim to manipulate the etiological underpinnings of disease and other behaviors. MacKinnon (2008) described the building of the Panama Canal example as one of the earliest applications of mediation analysis in this context, where researchers designed strategies to manipulate a mediator to ultimately impact an outcome. Although the French began building the canal in the late 1800s, their efforts halted after a yellow fever epidemic. The United States later picked up the project, but not without first leading a public health attack on yellow fever to protect workers from disease. Based on a theory that mosquitoes were carriers of yellow fever, doctors implemented measures to reduce workers' mosquito exposure (e.g., screened in sleeping quarters, improved drainage to reduce standing water). This

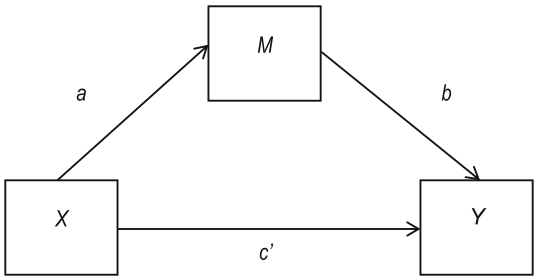

Fig. 23.1 Path diagram of the single mediator model. X = the independent variable (referred to as a program throughout this chapter), Y = the dependent variable, M = the mediator variable, a = the effect of the independent variable on the mediator, b = the effect of the mediator on the outcome controlling for X, and c' = the direct effect of X on Y controlling for M. The indirect effect of X on Y through M is captured by $(a)(b)$. In program evaluation, the a path of the model is referred to as the action theory of a program and illustrates how a program is designed to manipulate mechanisms of change; the b path of the model is referred to as the conceptual theory of a program and illustrates how the intervention intends to change the behavioral determinants

prevention effort was driven by the hypothesis that human exposure to mosquitoes mediated yellow fever occurrence. Thus, the intervention (i.e., the independent variable) sought to manipulate mosquito exposure (i.e., the mediator), with the intention of impacting yellow fever occurrence (i.e., the outcome).

Early Conceptualizations of Mediation in Social Science

Early work described mediators as process variables that contributed to understanding the underlying process of change in a program (e.g., Baron & Kenny, 1986; Judd & Kenny, 1981). One probed the underlying mechanisms of change with a series of causal steps to provide evidence for mediation. In the causal steps approach, one assessed (a) the overall relation between the program and an outcome, (b) the relation between the mediator and the program, and (c) the relation between the outcome and both the program and the mediator. The data supported mediation if the researcher found evidence for all three conditions, and there was a reduction in the magnitude of the overall relation between the program and the outcome once the mediator was modeled (Fig. 23.2).

Although Judd and Kenny (1981) originally advocated that a program must have a significant effect on the outcome to investigate mediation in the causal steps approach, they recognized that "if the mediating chain is a long one, the treatment may have a low correlation with the outcome" (p. 207). This acknowledgment marked the foundation of an argument now more widely accepted in the literature that the requirement of a significant overall effect of X on Y should be relaxed. The relaxation of this requirement is logical for several reasons. First, the test of the effect of X on Y is itself a statistical test and subject to statistical error just as any

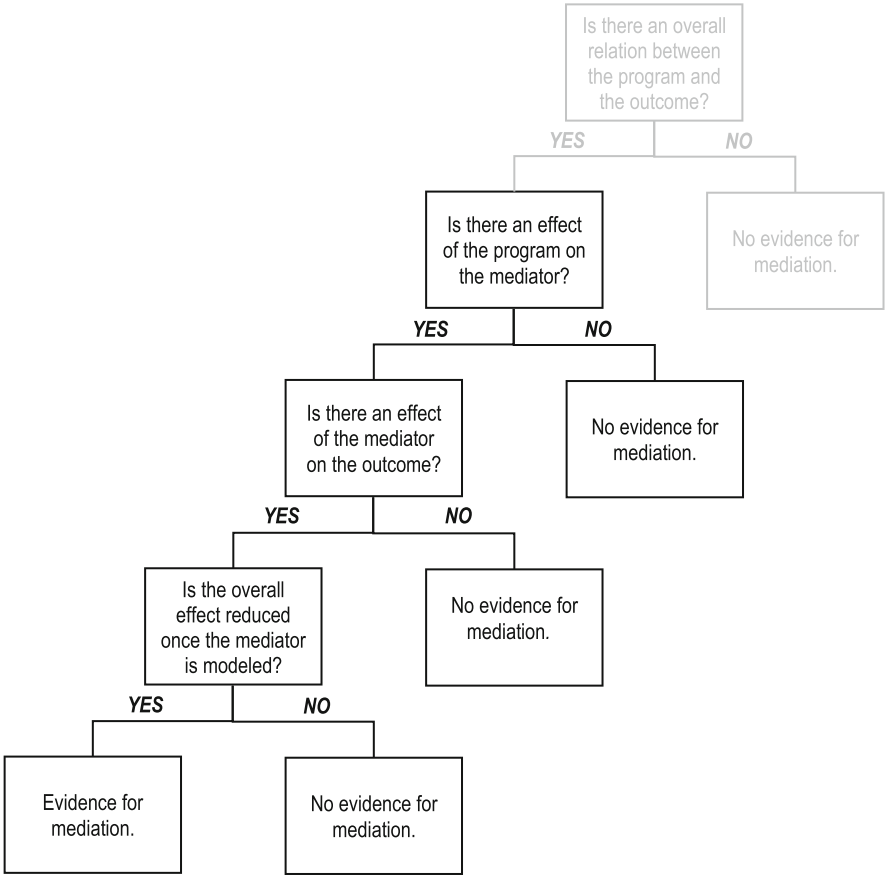

Fig. 23.2 Causal steps mediation procedure. The requirement that there be an overall effect of the program on the outcome has been indicated in gray text because this is widely accepted to no longer be a requirement for mediation

other. Thus, treating the test of *X* on *Y* as a population effect is fallible. Second, if the mediation process is hypothesized to be distal in nature (e.g., multiple mediators involved in a long causal chain between the independent and dependent variables), there will be limited power to test the overall effect of *X* on *Y*. Third, when the direct effect of *X* on *Y* and the indirect effect of *X* on *Y* through *M* have opposite signs in the mediation model, the overall effect of *X* on *Y* may be null despite significant mediation. Finally, and perhaps most critical to prevention work and program evaluation, investigating mediation effects in the presence of nonsignificant overall effects lends insight into why an intervention failed (see Section Mediation Analysis in Prevention Research for complete discussion). The information garnered from such an analysis can inform program refinement and can guide revision strategies in iterative program development.

Estimating Mediation

Two basic regression equations underlie the single mediator model shown in Fig. 23.1:

$$\hat{Y} = b_0 + c'X + bM, \tag{23.1}$$

and

$$\hat{M} = b_0 + aX, \tag{23.2}$$

where $\hat{Y}$ is the predicted value of Y, c' is the effect of the program (X) on the outcome controlling for the mediating variable (M), b is the effect of the mediating variable on the outcome controlling for the program, and a is the effect of the program on the mediator. $\hat{M}$ is the predicted value of M in (23.2). This chapter focuses on the context where the independent variable is a program, but the concepts and methods described here also apply to situations where mediation is investigated with different types of independent variables.

Although other tests of mediation are possible, contemporary methodological work indicates that the product of coefficients, $\hat{a}\,\hat{b}$, is the most flexible estimator of the mediated effect, extending well to more complex models (MacKinnon, 2008). Given decomposition rules of variable relations in path analysis, the mediated effect *ab* is actually the *indirect effect* of the program on the outcome transmitted through the intervening variable. The same logic defines the *direct effect* of the program on the outcome as c' and indicates that the *total effect* (i.e., the overall effect) of the program on the outcome, c, can be computed by summing the direct and indirect effect estimates from the model:

$$\hat{c} = \hat{a}\,\hat{b} + \hat{c}'. \tag{23.3}$$

Normal theory statistical tests based on the critical ratio of the mediated effect to its standard error are inaccurate for the mediation estimate because the product of two normally distributed variables is not itself normally distributed. Bootstrapped confidence intervals based on the empirical sampling distribution of $\hat{a}\,\hat{b}$ or asymmetric confidence limits based on the theoretical sampling distribution of the product *ab* are preferable for null hypothesis significance testing (MacKinnon, 2008). These more accurate methods are now widely accessible given the integration of bootstrapping options in several statistical software packages, as well as the creation of a program called PRODCLIN, which computes the distribution of the product test for mediation using user-specified values of $\hat{a}$ and $\hat{b}$ and their respective standard errors.

Effect Size Measures for Mediation

Effect size measures are an important addition to analysis when interpreting results, because they provide a concrete way to understand the practical, or clinical, significance of effects. The focus of effect size measures in mediation concentrates on the magnitude of different effects in the model (i.e., the indirect effect, the direct effect, and the total effect) to assess the relative contribution of each. Several different measures of effect size have been described for the mediated effect as well as for individual paths in the single mediator model (Fairchild, MacKinnon, Taborga, & Taylor, 2009; MacKinnon, 2008; Preacher & Kelley, 2011). Effect size measures for individual paths in the mediation model focus on the relation between two variables and are useful for understanding what pieces of the model work versus those parts that need to be improved. Effect size measures for the mediated effect quantify the magnitude of the mediated effect as a whole.

Effect size measures to describe individual paths in the mediation model include squared partial (or semi-partial) correlations and standardized regression coefficients. Squared partial correlations (or squared semi-partial correlations) quantify the amount of variance explained in the mediation model that is uniquely attributable to a single piece of the design, such as the variance in Y that is explained by M but not X or the variance in Y that is explained by X but not M. Standardized regression coefficients illustrate the strength of individual paths in the mediation model on a standardized metric. By identifying the relative contribution of individual paths in the mediation model, programs can be iteratively refined to promote stronger components and to remove or reduce weaker ones.

One commonly used effect size measure for the mediated effect is the proportion mediated. This measure represents the proportion of the total effect of X on Y that is mediated by the intervening variable (MacKinnon, Fairchild, & Fritz, 2007):

$$\frac{\hat{a}\,\hat{b}}{\left(\hat{a}\,\hat{b}\right)+\hat{c}'}. \tag{23.4}$$

The proportion mediated has also been recommended for the validation of surrogate endpoints in epidemiology and health-related research. A surrogate endpoint is a measure related to the study outcome in such a way that it can serve as a proxy for the outcome in analysis. Examples of surrogate endpoints are serum cholesterol levels for the ultimate outcome of coronary heart disease or the presence of polyps for the ultimate outcome of colon cancer. Despite being computationally easy, as well as being an intuitive measure, methodological work has shown that the proportion mediated only performs well with samples of greater than 500 (MacKinnon, Warsi, & Dwyer, 1995). This large sample size requirement may limit the utility of the measure given the prevalence of research with smaller sample size.

Another effect size measure for the mediated effect that can be used with smaller sample sizes is the R^2 mediated. The R^2 mediated quantifies the proportion of

variance in the outcome that is common to both X and M but is not attributable to either predictor alone (Fairchild et al., 2009). The R^2 mediated is computed by using squared bivariate correlations and the overall model R^2 from a model where Y is predicted from both X and M:

$$r^2_{MY} - \left(R^2_{Y,MX} - r^2_{XY}\right), \tag{23.5}$$

where r^2_{MY} is the squared correlation between the outcome and the mediator, $R^2_{Y,MX}$ is the overall model R^2 from the regression equation where Y is predicted from X and M, and r^2_{XY} is the squared correlation between the outcome and the independent variable (i.e., X). It is possible that estimates from the measure can be negative in some circumstances (a mathematical artifact of the equation). Because there are various advantages and disadvantages to different mediation effect size measures, it may be useful to use more than one to gain the clearest picture of effects in a model. Preacher and Kelley (2011) recently presented a new residual-based effect size measure for mediation that may handle some of the limitations with the measures described here, but the effect size measure described in that paper has not yet been statistically evaluated.

Assumptions of the Mediation Model

As with any statistical test, there are underlying assumptions of the mediation model that assure proper interpretation; violations of these assumptions may compromise the integrity of results. Assumptions of the model include basic ordinary least squares regression assumptions, correct causal ordering of the variables, no reverse causality effects, and no XM interaction. Nonrandomized studies make interpretation of causal relations in the mediation model tentative. Although the single mediator model is often represented in path analysis where causation is implied, causal effect estimation is not possible in some studies. The issue of causal inference in mediation has been, and continues to be, extensively studied. Given random assignment of X, the a path of the mediation model can be interpreted as a causal effect, but the b and c' paths of the mediation model are not causal parameters if the sample is not also randomly assigned to levels of the mediating variable (MacKinnon, 2008). In situations where the assignment mechanism to the mediator is not known, specific design strategies or statistical corrections may be implemented to strengthen causal inference. Design strategies that strengthen causal inference include, but are not limited to, within subject designs and matching subjects across treatment groups. Statistical corrections to strengthen causal inference include, but are not limited to, the use of propensity scores and more complex models such as principal stratification. The judicious selection of covariates is important for approximating subjective randomization in these situations.

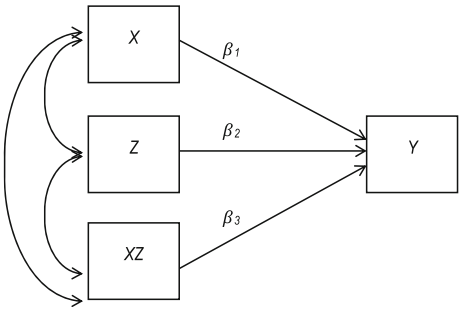

Fig. 23.3 Path diagram of the basic moderation model. X = the independent variable, Y = the dependent variable, Z = the moderator variable, XZ = the interaction of the independent and moderator variables, β_1 = the effect of the independent variable on the outcome controlling for Z and XZ, β_2 = the effect of the moderator on the outcome controlling for X and XZ, and β_3 = the effect of the XZ interaction on the outcome controlling for the lower order effects. The *double-headed curved lines* denote correlations among the predictor variables

The Moderation Model

What Is a Moderator?

A moderator is a third variable that interacts with the primary predictor variable to influence the outcome (Fig. 23.3). Thus, in moderation the regression of the predictor on the outcome differs across levels of the moderating variable. This dependency yields a nonadditive relation between the predictor and outcome variables, resulting in different regression relationships across varying levels of the moderator. Specifically, the moderator variable may affect the strength and/or direction of the relation between the predictor and outcome: enhancing, reducing, or changing the influence of the predictor. By definition, moderator effects yield differential prediction of the outcome by the primary predictor at different levels. The key difference between moderating and mediating variables is that mediators intercede in a causal sequence between two variables, whereas moderating variables do not.

The application of moderation is also particularly relevant in prevention research. One prominent moderation example in prevention work is the stress-buffering theory of social support in health psychology. Cohen and McKay (1984) and others have shown that social support has a protective effect against negative life events and stress, such that the relationship between negative life events and stress is lower for individuals with high levels of social support. In this example, social support acts as a moderator in the relationship between negative life events and ensuing stress, such that the relationship between the variables at different levels of social support differs.

Baseline-by-treatment interactions, where program effects are affected in either size or direction by baseline levels of the outcome variable, are another example of moderation in prevention research (Fritz et al., 2005). High-risk participants who start out lower on a positive outcome to begin with (or higher on a negative outcome to begin with) may get better gains from a program as they have more room to develop. Likewise, low-risk participants who start out higher on a positive outcome to begin with (or lower on a negative outcome) may not get as much benefit from an intervention. Flay and colleagues (1985) found that a program to decrease smoking behavior had a larger impact on individuals who already smoked themselves or who were exposed to smoking in their home as compared with those with no such exposure.

Estimating Moderation

The multiple regression equation for the basic moderation model is:

$$\hat{Y} = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ, \tag{23.6}$$

where $\hat{Y}$ is the predicted value of the outcome, β_1 is the effect of the program (X) on the outcome controlling for other variables in the model, β_2 is the effect of the moderator variable on the outcome controlling for other variables in the model, and β_3 is the effect of the interaction between the moderator and the program on the outcome. Again, although this chapter focuses on the context where the independent variable is a program, the concepts and methods described here extend to situations where the independent variable is of a different type.

The multiple regression framework permits evaluation of either categorical or continuous predictors, as well as higher order moderator effects such as three-way interactions or curvilinear interactions. A t-test of the β_3 coefficient in (23.6) tests the interaction term for statistical significance; if β_3 is significant then there is significant moderation in the data. Main effects do not accurately describe the relation between X and Y in the presence of moderation and should not be interpreted independently. Centering variables about their mean in analysis facilitates valuable interpretation of these lower order terms and reduces nonessential multicollinearity among predictors. Detailed rules for centering and the resultant interpretation of coefficients in the moderation model are provided in Aiken and West (1991).

Graphing simple slopes of Y on X at different values of Z elucidates the moderation relation. Rearranging terms in the basic moderation model facilitates plotting simple slopes of the interaction and illustrates how the slope of Y on X depends on different values of the moderator (Aiken & West, 1991):

$$\hat{Y} = (\beta_1 + \beta_3 Z)X + \beta_2 Z + \beta_0. \tag{23.7}$$

When plotting interactions with categorical moderator variables, values of Z should correspond to the groups that you wish to plot. For continuous moderator variables, theory or other important criteria may provide meaningful values of Z at which to plot the X–Y relation. Otherwise, simple slopes may be evaluated for scores at ± 1 standard deviation from the mean of Z and at the mean of Z itself, the rationale being that these values represent a range in which most observed scores should fall in a normal distribution.

Effect Sizes Measures for Moderation

Squared partial correlation coefficients can guide understanding of the practical significance of effects in the moderation model. As described earlier, r^2 measures illustrate the unique portion of variance explained by a predictor in the outcome that is not explained by other predictors in the model (Cohen, Cohen, West, & Aiken, 2003). By evaluating an R^2 change statistic ($R^2\Delta$), one may also quantify variance accounted for in the model by the interaction above and beyond that which is explained by the main effects:

$$F = \frac{\left(R_2^2 - R_1^2\right) \Big/ \left(k_2 - k_1\right)}{\left(1 - R_2^2\right) \Big/ \left(N - k_2 - 1\right)}, \tag{23.8}$$

where R_2^2 is the R^2 associated with a model that includes both the main effects and the interaction term, R_1^2 is the R^2 associated with a regression model that includes only the main effects for X and Z, k_2 is the number of predictors in the equation that includes the interaction term, k_1 is the number of predictors in the equation that includes only the main effects, and N is the sample size.

Assumptions of the Moderation Model

Several statistical assumptions underlie the moderation model, including those assumptions associated with ordinary least squares regression (Cohen et al., 2003). An additional assumption of the moderation model is homoscedasticity. Homoscedascity requires that the residual variance that remains after predicting Y from X be constant across values of the moderator variable. Bartlett's M statistic provides a formal statistical test of whether homogeneity of error variance has been violated (nonsignificant results indicate that no violation has occurred). Nonparametric tests (e.g., Kruskal-Wallis) are available in the event that the homoscedasticity assumption is violated.

Using Mediation and Moderation to Achieve Prevention Research Goals

The utility of mediation and moderation analyses in prevention research is multifold and their integration into analysis is a critical component of program evaluation. If the major goals of prevention research are to successfully implement interventions to make an impact as well as to increase our scientific understanding of behaviors and prevention, we can benefit from moving beyond only evaluating overall program effects to also investigating mediating agents of behavioral change and possible moderating influences on program efficacy. The mediation and moderation models broaden the scope about what we can learn from our interventions and facilitate their efficacy.

Mediation Analysis in Prevention Research

Mediation analysis has the potential to inform more than just how a program achieved its effects, or alternatively why it did not achieve effects. By integrating mediation analysis into program evaluation, a researcher can conduct a manipulation check of a prevention or intervention program. Specifically, one can statistically test whether a program successfully manipulated the purported behavioral determinant(s) and thus determine whether the program changed what it was designed to change. For example, if a program intended to promote healthful eating by increasing social norms for healthful eating, then program effects should be observed on social norms for healthful eating. The absence of such a finding would indicate that the program failed to manipulate what it intended. Conducting a mediation analysis also permits the identification of variables that require improved measurement or additional manipulation in a program, as formalizing the measurement of mediating constructs enables explicit investigations of their reliability and validity. Such identification is especially important because poor measurement of variables has the propensity to obscure treatment effects, making otherwise viable prevention efforts appear ineffective.

Prevention efforts can be iteratively refined to be more efficient and cost effective by identifying successful and unsuccessful components of a program. Whether a program has been successful or ineffective, mediation analysis can identify components of the manipulation that contributed to its success or failure, illuminating strengths as well as limitations. By evaluating specific components of a program that produced intended or unintended change, mediation analysis pinpoints (a) *supportive elements* or those components that encouraged intended behavior, (b) *inert elements* or those components that did not contribute to changing the behavioral outcome, and/or (c) *iatrogenic elements* or those components that promoted unintended negative effects of the program. The differentiation of these components streamlines prevention and intervention efforts by facilitating

concentration on effective mechanisms of behavioral change; such concentration is paramount in resource-limited circumstances. Beyond this, mediation analysis also has the potential to direct the exploration of competing theories for behavior as researchers can explicitly test and analyze alternative mediators together in multiple mediator models (MacKinnon, 2008; MacKinnon & Dwyer, 1993). MacKinnon (2008) details how to estimate and evaluate these types of models.

Mediation analyses may be framed as *mediation for explanation* or *mediation by design* (MacKinnon, 2008). Mediation for explanation approaches involve probing the underlying mechanisms of an effect after is has occurred. For example, research has shown that childhood physical abuse relates to patterns of violence later in adulthood. Dodge, Bates, and Pettit (1990) informed this relation by showing evidence that physically abused children acquire deviant patterns of processing social information (i.e., a mediator), which in turn predict aggressive behavior in adulthood. Although this approach is vulnerable to capitalizing on chance findings in the data, it is useful in that it forces consideration of how an effect occurred. To that end, findings from a mediation for explanation study may help guide future mediation by design studies.

Mediation by design approaches identify mediator variables a priori and explicitly design an intervention to manipulate that mediating variable with the intention of ultimately impacting an outcome. In this context, theory defines programmatic components that ostensibly relate to an outcome, and successful manipulation of these determinants should change the target behavior. The Panama Canal example presented at the beginning of this chapter represents a mediation by design application. A second example of mediation by design is John Snow's intervention in the mid-nineteenth century London cholera epidemic. Snow, a physician, theorized that the disease was being propagated by a contaminated water pump in the city and had officials shut it down. By removing the contaminated water supply, and thus manipulating the mediator of disease, Snow was able to interrupt the cholera outbreak. The mediation by design approach is the foundation of instigating behavioral change in an intervention context and embodies the heart of prevention science. If putative mediators are causally related to outcomes as theorized, then a program designed to change the mediators should lead to change in the outcome.

Component paths of the mediation model have distinct interpretations in a program evaluation context. The *action theory* of a program is defined by the path relating the program to the mediator (i.e., the *a* path in Fig. 23.1) and illustrates what components of the program are designed to manipulate mechanisms of change (Chen, 2005; MacKinnon, 2008). Action theory explicitly outlines pieces of the program curriculum related to hypothesized mediator(s) and illustrates how the intervention intends to change those behavioral determinants. The *conceptual theory* of a program is defined by the path relating the mediator to the outcome (i.e., the *b* path in Fig. 23.1) and describes how hypothesized mediators are theoretically related to outcomes of interest (Chen, 2005; MacKinnon, 2008). This link thus defines the theory that connects variables or psychological constructs to behavioral outcomes and quantifies the mediator's ability to affect the outcome.

Relationships examined in this part of the model are driven by previous research or theory that explains antecedents of behavioral outcomes.

Evaluation of action and conceptual theory components in the mediation model can direct future actions of program stakeholders in the face of program failure. For instance, if program failure was driven by an inability to manipulate mechanisms of change (i.e., a failure in action theory), it may be sensible to move forward with the approach but to explore additional resources or alternative design strategies to enhance the manipulation in later implementations. Alternatively, if program failure was driven by the lack of a relationship between the mediator and the outcome (i.e., a breakdown in conceptual theory), it may be reasonable to review the original theoretical conceptualization of the model and reconsider underlying determinants of the target behavior. Although the failure to find effects of a mediator on the outcome often indicates that the variable may not be instrumental in instigating change, such a finding may also point toward sleeper mediation effects or effects that do not manifest until a later time (MacKinnon, 2008). Sleeper mediation effects may be observed with mediators that take time to have an impact, such as lowering cholesterol on reducing the incidence of cardiovascular disease.

The effectiveness of any given program hinges on the quality of both action and conceptual theories, and understanding the success or failure of each piece guides future steps in program development. If a mediation analysis indicates failure in the action theory of a program, stakeholders should attempt to identify elements that affect effective program implementation and delivery. In this context, researchers may focus on resource development or look toward improving implementation quality to enhance the program's ability to manipulate the mediator and to intensify treatment effects. It may also be valuable to reconsider whether the mediator can be experimentally manipulated and if so what factors need to be considered to maximize effects. In contrast, if the mediation analysis indicates a failure in the conceptual theory of a program, stakeholders should reassess their assumptions of what variables determine the outcome and should reevaluate the theory underlying relations between variables in the model. Finally, if a mediation analysis reveals failure in both the action and the conceptual theories, it may be important to consider whether the program is worthwhile to pursue (Chen, 2005).

Moderation Analysis in Prevention Research

Investigating moderator variables is an important aspect of streamlining and refining prevention and intervention programs, because contextual effects can enhance or detract from the efficacy of program components. Moreover, the presence or absence of moderation effects in a theory-based design characterizes the generalizability, or the external validity, of an intervention. Investigating moderator variables can identify participants for whom the intervention is most effective, can compare relative effectiveness of different delivery or implementation strategies, and can identify characteristics of the environment that may exert differential

influences on program effects. Such investigations may be especially important in the presence of unexpectedly small or nonsignificant program effects, where overall treatment effects can be null because there are different subgroups for which the program operates in different ways. To enable these investigations, moderator variables should be integrated into measurement batteries at the beginning of a project. Contingent on their role, different program stakeholders may have different research questions regarding moderation of intervention effects. One may be more or less interested in whether a program effectively reaches a target group, whether service providers are effectively implementing the intervention, or whether contextual factors affect effective delivery of the intervention. Factors that enhance or detract from treatment fidelity among participants and the broader implications of this on public policy may also be relevant.

Moderator variables can be continuous or categorical, as well as mutable or constant, and may stem from a variety of different sources. As described earlier, baseline measures of outcome variables may moderate program effects by driving baseline-by-treatment interactions. Different facets of the environment in which the program is implemented can affect program effects as well, such as type of school (i.e., public vs. private) or geographical location. Finally, characteristics of the participant/service provider interactions and/or treatment fidelity may also moderate program effects. It is useful to consider Bronfenbrenner's Ecological Systems Theory (1979) to put these varied sources into context. Bronfenbrenner suggests that an individual's development is directed by a multifaceted, nested environmental system whose levels affect one another bidirectionally. Central to this system is the individual, whose personal characteristics such as age, gender, genetic phenotypes, or baseline levels of the outcome may moderate program impact. Proximal factors just outside the individual, called microsystems (e.g., family, school, or home), can also act as moderators. An example of a moderator from this part of the system could be exposure to deviant peers, which may have the potential to alter program effects. Exosystems, or the social influences beyond an individual's proximal contexts such as the larger neighborhood in which he or she lives or the mass media to which he or she is exposed, also have the potential to moderate program effects. Some examples of moderators from this part of the framework might be exposure to a mass media campaign related to the program outcome or availability of different community resources that may affect structural social support. Macrosystems, or the larger cultural context to which an individual is exposed, can generate moderating influences as well. Some examples of moderators from this part of the framework include ethnicity, socioeconomic status, or prevailing political ideology. Finally, influences from the chronosystem, or the broader temporal context representing dynamic changes in personal and environmental influences over time, may also moderate program effects. A potential moderating influence from this aspect of the model could be something like time since divorce, as disruptive effects for children have been noted to be higher in certain timeframes.

Some moderator variables may be more or less amenable to program tailoring. Although program subgroups may be formed on hypothesized moderators such as

age or gender with little difficulty, forming program subgroups based on other moderator variables may be impractical and/or unethical in some circumstances. Further it may actually be iatrogenic to select persons based on risk-taking behavior. For example, Dishon, McCord, and Poulin (1999) found that grouping persons at risk for drinking in an intervention led to more alcohol consumption for that group, possibly because of a pro-alcohol norm within the subset of high-risk individuals. Nonetheless, the identification of subgroups for which a program is most effective remains useful both for efficient program development and for theory regarding the process by which programs work across different subgroups.

Simultaneously Analyzing Mediation and Moderation in Program Evaluation

Simultaneously estimating the mediation and moderation models provides the opportunity to examine interesting effects not possible in either model alone. For example, it may be of interest to investigate whether the mediation relation holds across different groups of participants (e.g., men vs. women or low risk vs. high risk) to determine the generalizability of a mediating mechanism or to explain an unexpectedly small mediated effect. To test such a hypothesis, a researcher evaluates whether *the indirect effect is moderated*. Several contemporary methodological resources lay out statistical strategies for evaluating the moderation of indirect effects (e.g., Edwards & Lambert, 2007; Fairchild & MacKinnon, 2009). With continuous moderators, the researcher can estimate a single mediation model by incorporating applicable interaction terms (MacKinnon, 2008). Significant interaction effects in either the *a* or the *b* path of the model imply moderation of the indirect effect across some range of the continuous moderator (Fairchild et al., 2009). With categorical moderators, this single model approach can be applied as a combined-groups model. Alternatively, the researcher can apply an individual-groups model by estimating separate mediation models at each level of the categorical moderator variable and testing parameters of interest for equivalency across the subgroups. Analyzing simple mediation effects at different levels of the moderator variable is an important aspect of these analyses. Investigating simple direct effects in the mediation model may also be valuable depending on the scope of the relevant research questions.

The combination of mediation and moderation analyses also permits the investigation of whether an interaction effect can be explained, at least in part, by a mediating mechanism. To explore this *mediation of a moderator effect* scenario, one evaluates whether the interaction between the independent variable and the moderator predicts a mediating variable, which in turn predicts the outcome. A test of moderation in the *a* path of the mediation model thus suffices as a test for this effect. Fairchild and MacKinnon (2009) provided a mediation of a moderator effect example where part-time versus full-time work status moderated outcomes of a

worksite wellness program. A mediation model was able to explain this effect such that full-time workers were getting more exposure to program-related social norms at the work place, contributing to their larger program effect.

Combined models for mediation and moderation often have been a source of confusion for substantive researchers. This is not surprising given the equivalence of some effects across the models. Focusing on which effects (i.e., overall vs. direct vs. indirect) and/or which paths in the mediation model are moderated avoids this confusion while encouraging the researcher to focus on specific research hypotheses (Edwards & Lambert, 2007). This focused approach may be especially valuable in program evaluation work as it facilitates explicit examination of heterogeneity in program action and conceptual theories to explore whether a program impacts behavioral determinants in the same way across different groups of participants or to ask whether the mediator is related to the outcome in the same way across different groups of participants, respectively.

Concluding Remarks and Recommendations

Intervention and prevention programs are best conceptualized as multifaceted efforts that link scientific theory and process evaluation. Program development and evaluation within this context should not be limited to assessing the efficacy of programs and/or the tenability of theory but should also investigate how to improve service delivery and dissemination to target participants in evidence-based ways. Evaluating potential program mediators and moderators contributes to this goal and promotes accountability and thoughtful curriculum development based on internal and external validity of program components. Researchers can ameliorate prevention programming by using mediation and moderation analyses to iteratively refine interventions in an open system where program theory and empirical application are considered in tandem. We provide a general program development strategy that incorporates mediation and moderation analyses to enable these efforts (Fig. 23.4).

Mediation analysis fosters effective program development and refinement by enabling investigations of how or why a program was successful (or unsuccessful) in changing target behaviors. By differentiating supportive versus inert intervention components, program stakeholders can streamline curricula and reduce intervention costs. It may also be possible to garner a broader look at program fidelity and implementation issues by looking at serial mediation models. For example, one might posit that the efficacy of implementation directly affects a program, which then affects the agent of change, which is hypothesized to affect the outcome of interest (Fig. 23.5).

Parsing out the relationships in this way can help further identify strengths and/or weaknesses in a program and where to concentrate efforts on how to fix them. Moderation analysis also fosters program refinement by enabling program stakeholders to examine the generalizability of findings across different groups or settings, as well as to identify variables that may improve or reduce program

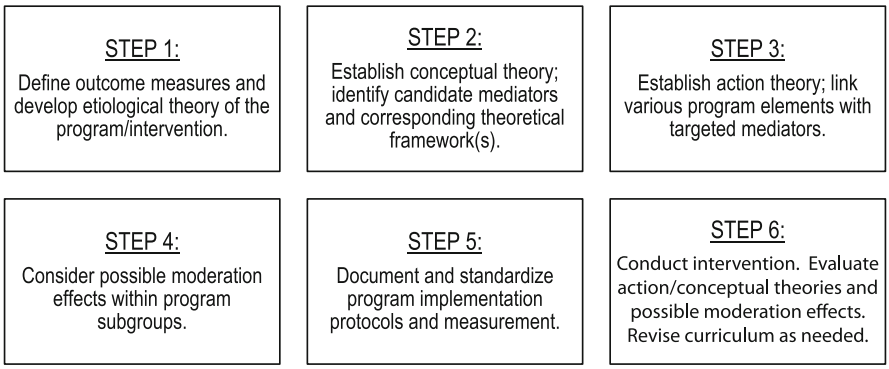

Fig. 23.4 A mediation- and moderation-based program development strategy. Figure adapted from a table originally presented in MacKinnon (2008, p. 42)

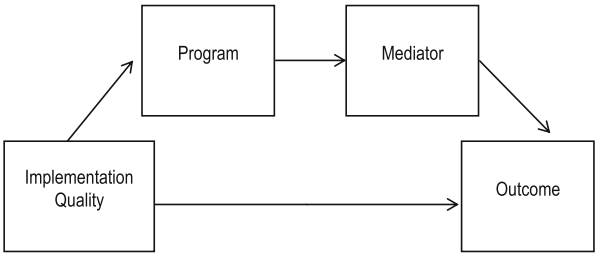

Fig. 23.5 Investigating implementation efficacy: a serial mediation model

efficacy. The evaluation of moderator effects thus enables tailoring interventions by identifying groups of individuals that may need more or fewer services or perhaps none at all. Targeting interventions in this manner streamlines limited resources.

Mediation and moderation analyses have the ability to impact and encourage all levels of prevention efforts: (a) universal efforts where individuals have not been identified on the basis of individual risk that intend to benefit all members of a group, (b) selective efforts that focus on risk reduction and target a subset of individuals who have been determined to be at high risk for developing a problem, and (c) indicated efforts that focus on harm reduction and target individuals who have detectable signs or symptoms of a particular problem but do not currently meet diagnostic criteria. Commitment to a dynamic process of continual evaluation and development is critical to maximize our ability to affect behavioral change across all of these levels. Ultimately, understanding moderators of program dissemination and implementation, as well as identifying mediating mechanisms of behavioral change can expand the theoretical body of knowledge underlying prevention programs and social issues more broadly.

Acknowledgement Work reflected in this chapter was supported in part by the National Institute on Drug Abuse, under grant R01DA030349 awarded to the first author.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publication.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Chen, H. (2005). *Practical program evaluation: Assessing and improving planning, implementation, and effectiveness*. Thousand Oaks, CA: Sage Publications.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. A. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, S., & McKay, G. (1984). Social support, stress, and the buffering hypothesis: A theoretical analysis. In A. Baum, S. E. Taylor, & J. E. Singer (Eds.), *Handbook of psychology and health* (pp. 253–267). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dishon, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755–764.

Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990). Mechanisms in the cycle of violence. *Science, 250*, 1678–1683.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods, 12*, 1–22.

Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science, 10*, 87–99.

Fairchild, A. J., MacKinnon, D. P., Taborga, M., & Taylor, A. B. (2009). R2 effect-size measures for mediation analysis. *Behavioral Research Methods, 41*, 486–498.

Flay, B. R., Ryan, K. B., Best, J. A., Brown, K. S., Kersell, M. W., d'Avernas, J. R., et al. (1985). Are social-psychological smoking prevention programs effective? The Waterloo Study. *Journal of Behavioral Medicine, 8*, 37–59.

Fritz, M. S., MacKinnon, D. P., Williams, J., Goldberg, L., Moe, E. L., & Elliot, D. L. (2005). Analysis of baseline by treatment interactions in a drug prevention and health promotion program for high school male athletes. *Addictive Behaviors, 30*, 1001–1005.

Judd, C. M., & Kenny, D. A. (1981). *Estimating the effects of social interventions*. New York, NY: Cambridge University Press.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17*(2), 144–158. doi:10.1177/0193841X9301700202.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593–614.

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research, 30*, 41–62.

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods, 16*, 93–115.

Saunders, D. R. (1956). Moderator variables in prediction. *Educational and Psychological Measurement, 16*, 209–222.

Wholey, J. S. (1987). Evaluability assessment: Developing program theory. In L. Bickman (Ed.), *Using program theory in evaluation* (New directions for program evaluation, Vol. 33). San Francisco, CA: Jossey-Bass.

Chapter 24
Advances in Missing Data Models and Fidelity Issues of Implementing These Methods in Prevention Science

Shelley A. Blozis

Introduction

Data analysis with incomplete data is common in prevention research (Graham, Hofer, Donaldson, MacKinnon, & Schafer, 1997). Appropriate applications of methods to treat missing data are essential in ensuring unbiased parameter estimates, as well as in maintaining optimal efficiency in parameter estimates. The primary challenge for researchers if some data are missing concerns the particular analytic method that is deemed necessary for a data analysis, such as ANOVA, and the source or mechanism that gave rise to the missing data. That is, different analytic methods make different requirements of the data (e.g., analysis of variance requires complete data), and along with these requirements are specific assumptions about the source of the missing data. Careful planning of research studies when missing data are anticipated can greatly reduce the adverse effects of missing data and improve statistical inference. Knowledge of methods for handling missing data is therefore critical in the planning stages of a prevention study (Graham & Donaldson, 1993).

Missingness

In the missing data literature, whether *missingness* (i.e., whether a response is missing) is ignorable or nonignorable given a particular plan for data analysis is central to missing data problems. For a given data analysis, missingness is ignorable if the parameters of a data model (i.e., a model of the hypothesized relationships between the variables of interest) are independent of the mechanism that generated

S.A. Blozis (✉)
Department of Psychology, University of California, Davis, CA, USA
e-mail: sablozis@ucdavis.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_24,

the missing data. In other words, the missingness for a given problem is ignorable if the parameters of a data model are essentially the same whether or not the source of the missing data is taken into account. Thus, if the missing data mechanism is ignorable, the parameter estimates of a data model may be interpreted without concern for bias due to the missing data. Missingness is not ignorable if valid interpretation of a data model depends on the source of the missingness. Indeed, if the missingness is not ignorable and the missing data process is not addressed in the data analysis, parameter estimates may be biased. Several strategies have been developed to handle missingness that is not ignorable. To understand the role of missingness in a data analysis more generally, it is useful to first describe the types of data that arise when some data are missing. These types are described next.

Data Types

Little and Rubin (2002) provide a classification for data that are incomplete. For a given data set, the data are assumed to include both observed and missing values. The classification scheme depends on the relationships between the data, both the observed and the missing values, and whether data are missing or not. Thus, in addition to the data, one or more indicator variables are created to represent the missingness and are included in the data set. An indicator of missingness may, for instance, represent a particular pattern of missing data, such as whether an individual has completed participation in a prevention study, or a separate indicator variable may be created for each variable that has missing responses. A missingness indicator variable is usually set equal to 1 if an individual has a particular pattern of missing data and set equal to 0 otherwise. Different indicator variables may be created to represent the various missing data patterns that describe a particular data set (Schafer & Graham, 2002; also see Hedeker & Gibbons, 1997).

The first type of data is data that are missing completely at random (MCAR). Data are MCAR if the missingness is independent of both the observed and the missing data. In a prevention research study, data are MCAR if, for example, whether data are missing or not is independent of both the observed and the missing data. Examples of data that are MCAR are those for which missing data are planned as part of the data collection procedure, such as when participants are selected at random to receive one of multiple possible subsets of the study variables or to be measured according to one of several possible planned sets of measurement occasions of a longitudinal study. This type of missing data is referred to as *planned missingness* and may be used to reduce study costs or participant fatigue (Graham, Taylor, Olchowski, & Cumsille, 2006). If data are MCAR, then there are no systematic differences in the observed and missing values between those individuals with complete data and those with a pattern of missing data. Statistical inference from analytic methods that require complete data, such as ANOVA, is valid when data are MCAR. That is, if data are MCAR, complete-data methods yield unbiased estimates of parameters, although such approaches result in reduced sample sizes and statistical power.

The second type of data is data that are missing at random (MAR). Data are MAR if the missingness is independent of the missing data but is dependent on the observed data. More specifically, the missingness is not related to the missing data after accounting for relationships between the missingness and the observed data. In a prevention research study, for instance, in which measures of substance use are missing for some individuals, data are MAR if the likeliness that an individual will respond to an intervention survey concerning substance use is related to any variable other than the missing substance use outcome. Data are MAR if, for instance, older participants are less likely to respond to the substance use survey. In this example, the missingness depends on the age of the participant and not on the missing substance use data. More generally, the missingness in a given study may be related to one or more of the observed variables included in a data model.

Methods that are valid when data are MCAR or MAR are then suitable choices for a data analysis. Such methods include structural equation models and mixed-effects models that rely on maximum likelihood (ML) estimation procedures and make use of all of the observed data. In a longitudinal study, for instance, data are MAR if the missingness is related to the observed response measured prior to dropout but, conditional on this relationship, the missingness cannot be related to the missing data at the time of dropout or thereafter. Some types of planned missingness may also generate situations in which data are MAR (as opposed to MCAR), such as if study participants are selected for follow-up assessments based on their scores at a previous occasion, such as a pretest score. Here, the missingness is related to the observed pretest score, and conditioning on that relationship, may not be related to the missing score.

The third type of data is data that are not missing at random (MNAR). Data are MNAR when the missingness is related to the missing data even after taking into account the relationships between the missingness and the observed data. In a prevention study, data are MNAR if individuals with high levels of substance use are less likely to report their levels of use. In this case, the missingness depends on the missing substance use data. If data are MNAR, an analysis that is based on the assumption that data are MCAR or MAR may result in biased parameter estimates that may then lead to wrongful conclusions about the relationships among the variables of a data model. Thus, data that are MNAR present a missing data problem that is not ignorable. Indeed, the consequences of ignoring data that are MNAR are a source of great concern for researchers (Molenberghs & Kenward, 2007).

Data Analysis with Missing Data

The need to address missing data in a data analysis has resulted in a variety of approaches to missing data, including analysis that is based on only cases with complete data, in addition to a variety of imputation techniques that aim to generate complete data so that complete-case procedures may be applied. Imputation

methods include both single and multiple imputation techniques. Understanding the requirements and possible consequences of the different methods is essential to making informed decisions about a missing data method. Schafer and Graham (2002) provide a comprehensive review of many ad hoc missing data methods and consequences of their applications. Here, imputation methods and likelihood-based methods are reviewed.

Single imputation methods are those that aim to replace missing data with a single value. Common approaches include mean substitution in which missing values are replaced by the mean of the observed scores and regression substitution in which missing scores are replaced by values predicted by the observed data. Once more commonplace in the treatment of missing data, single imputation methods are generally avoided due to the serious problems associated with them, including biased parameter estimates and errors in statistical inference. Importantly, these ad hoc methods do not address the uncertainty involved in replacing the missing data. That is, imputed values represent estimates of the missing values, but ad hoc methods were not designed to address this uncertainty. A consequence of ignoring this uncertainty is that the standard errors of related parameter estimates tend to be underestimated, consequently leading to an increase in the type I error rate.

Multiple imputation (MI) is a technique in which missing data are replaced by multiple imputed values (Rubin, 1978, 1987). A common implementation of MI is by using a computer simulation process in which a set of observed data provide information about the relationships among a set of variables that form the basis of an imputation model. Missing values are then replaced by multiple random draws from the assumed distributions of the variables of the imputation model. The result is a set of multiple imputed data sets that are then analyzed individually using a complete-case analysis procedure. Parameter estimates obtained from each analysis are averaged across the set of results to obtain a single set of estimates. Standard errors of the estimates are generated to take into account variation in the estimates both between and within the imputed data sets. Thus, the standard errors take into account the added variation in the parameter estimates that is due to the imputation procedure. It should be noted that under MI the estimated standard errors of the parameter estimates will generally be greater than those that would have been produced had the data set been complete.

Not all of the variables that are included in an imputation model need to be included in a data analysis, but all of the variables to be included in a data analysis should be included in the imputation model. That said, MI requires that the data used for the imputation process are MCAR or MAR. Finally, it is also important to note that MI is model based. That is, MI is done under assumptions made about the predictive distributions of the missing data. Consequently, a different imputation model may result in a different set of imputed data.

Likelihood-based methods offer an alternative to complete-data methods and, similar to MI procedures, relax the assumption of MCAR. Full information maximum likelihood (FIML), a type of ML estimation that relies on raw data rather than sufficient statistics for the data analysis, may be used to estimate structural equation

models and mixed-effects models, for instance, and provide unbiased estimates when data are MCAR or MAR. By making use of all observed data, these methods do not suffer from a reduction in the precision of parameter estimates and statistical power, problems that have been well documented for analytic approaches that rely on only complete cases.

Under certain conditions, analyses based on MI and FIML have been shown to yield similar results (Schafer, 1997). If the two methods are applied to the same set of variables and a sufficiently large number of imputed data sets are generated when using MI, the two methods yield consistent results. Both methods have been shown to yield superior results relative to single imputation methods because these methods provide a better representation of the data (see Little & Rubin, 2002). Multivariate analysis using FIML or MI procedures is possible using major software programs, such as SAS. Several more specialized software programs have also made these procedures available, including those designed for the estimation of structural equation models, such as LISREL (Jöreskog & Sörbom, 2006) and Mplus (Muthén & Muthén, 1998–2010).

Ignorable and Nonignorable Missing Data

As discussed earlier, a central issue in performing an analysis with missing data concerns whether or not the missingness is ignorable. In summary, for complete-data methods, missingness is ignorable if data are MCAR but not MAR or MNAR. For MI and likelihood-based methods, missingness is ignorable if data are MCAR or MAR but not if data are MNAR. Two key issues arise from this information: The first concerns how to handle a data analysis when the missing data process is not ignorable. The second issue concerns how to evaluate assumptions of missingness because all three of the data types (MCAR, MAR, and MNAR) involve an assumption about the relationship between the missingness and the missing data, a relationship that cannot be empirically tested. That is, given that the missing data are not available for study, it is not possible to evaluate the relationships between the missingness and the missing data.

Data Analytic Approaches to Nonignorable Missingness

Several major frameworks have been developed for the analysis of data that are MNAR. Generally, these may be grouped into those that treat the missing data prior to the data analysis and those that handle the missing data and the data model simultaneously. MI procedures fall under the first type and may be used to treat data that are MNAR by incorporating information from correlates of the data, referred to as auxiliary variables, into the imputation process. This strategy may then lead to an ignorable missing data problem for the data analysis. Methods that fall under the

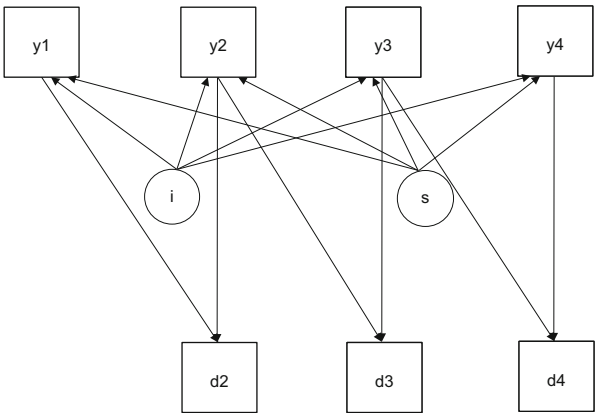

Fig. 24.1 Nonignorable missingness in a selection model

second type involve either an explicit model for the missing data process or an extension of a data model to directly include correlates of the data.

MI with Auxiliary Variables. MI may be used to address data that are MNAR by including correlates of the data in the imputation process. If auxiliary variables provide the information needed about the data, then the data are what is referred to as auxiliary variable MAR (or A-MAR) and the missing data process becomes ignorable (Daniels & Hogan, 2008). Data analysis may then be performed on the imputed data using only the variables of the data model. Using MI with auxiliary variables can provide superior results to MI methods that do not take advantage of these variables (Rubin, 1997).

MNAR Models. Nonignorable missingness may be addressed directly by extending a data model to include a model for the missing data process. Such models include selection models, shared-parameter models, and pattern-mixture models. In a selection model, the missingness depends on the data. In a longitudinal study, for instance, a selection model may be specified in which the missingness depends on the missing response at the time of dropout and the observed responses prior to dropout, as shown in Fig. 24.1 (Diggle & Kenward, 1994). In Fig. 24.1, indicators of missingness (e.g., dropout status) at occasions 2, 3, and 4, denoted by the variables *d2*, *d3*, and *d4*, depend on the observed response at the previous occasion (assuming an individual has not dropped from the study by that point) and the current measure that is missing if the individual has dropped or observed if the individual remains in the study by that occasion.

In a shared-parameter model, a special case of a selection model, the missingness depends on the random coefficients of a data model. The missingness may, for example, depend on the random intercept and slope of a mixed-effects model that is used to characterize longitudinal data, as depicted in Fig. 24.2 (Wu & Carroll, 1988). In Fig. 24.2, indicators of missingness at occasions 2, 3, and 4 are denoted by the variables *d2*, *d3*, and *d4*. Each indicator of missingness is specified to depend on the random intercept and random slope, represented by the variables

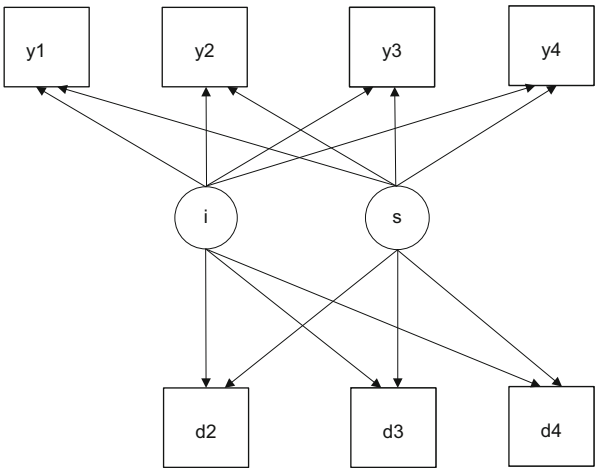

Fig. 24.2 Nonignorable missingness in a shared-parameter model

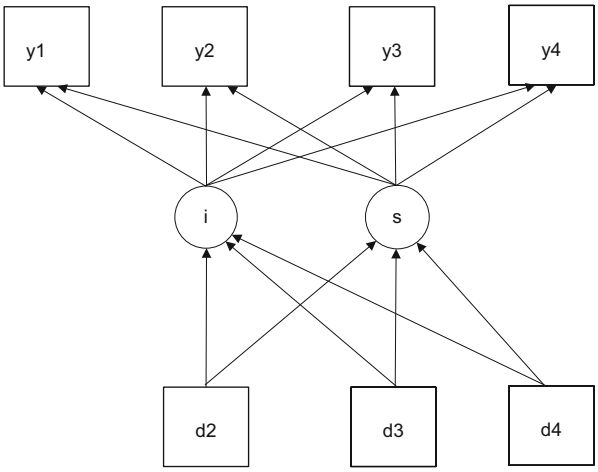

Fig. 24.3 Nonignorable missingness in a pattern-mixture model

i and *s*, respectively. In this example, the missingness depends on the expected value of an individual's response at a given measurement occasion, such as at the start of a prevention study, as well as on an individual's expected rate of change in the response over the study period.

In a pattern-mixture model, the response is specified to depend on the missingness. In a pattern-mixture random-effects model, for instance, the random intercept and slope of a model for a longitudinal response may depend on an indicator of missingness, such as an indicator of dropout, as shown in Fig. 24.3 (Hedeker & Gibbons, 1997). In Fig. 24.3, indicators of missingness at occasions

2, 3, and 4, denoted by the variables *d2*, *d3*, and *d4*, moderate the random intercept and random slope, represented by the variables *i* and *s*, respectively. In this model, an individual's expected response status at a given measurement occasion and expected rate of change in the response over time may be moderated by the dropout status.

Unlike MI methods, the MNAR models described require an explicit model for the missing data process. That is, the relationship between the missingness and the observed and missing data must be explicitly defined. This can present a challenge to researchers because the specific mechanism that gave rise to the missing data may not be known. Thus, it is often recommended that multiple missing data models, theoretically driven, be specified and fitted to the data so that the analysis does not rely on any single missing data model. In this way, competing ideas about a missing data process may be evaluated.

Models that Include Auxiliary Variables. Data that are MNAR may also be handled by fitting a data model that is extended to include correlates of the data. That is, the data model is extended to allow one or more auxiliary variables to correlate with variables of the data model. In this way, these models do not require an explicit model for the missing data process as required by selection, shared-parameter, and pattern-mixture models, and the data model is not altered with regard to the hypothesized relationships that define the data model. Estimation of models that include auxiliary variables may be carried out using FIML. Similar to MI implemented with auxiliary variables, information about the missing data is drawn from the auxiliary variables. Unlike MI, however, adding auxiliary variables as correlates into a data model does not require the generation of multiple data sets that must then be analyzed and summarized through additional steps.

Graham (2003) describes a saturated correlates model in which auxiliary variables are allowed to correlate with the exogenous variables, the residuals of the manifest dependent variables, the residuals of any indicators of latent variables, and each other. An example of a saturated correlates model applied to a longitudinal measure is shown in Fig. 24.4. In Fig. 24.4, a set of auxiliary variables are shown to correlate with the residuals that result from the regressions of the four measured responses (Y_1,. . ., Y_4) on the random intercept (I) and slope (S), as well as with the random intercept and slope that represent the exogenous variables of the longitudinal data model.

Collins, Schafer, and Kam (2001) conducted a data simulation study to better understand the role of auxiliary variables in a data analysis when data are incomplete. In their study, auxiliary variables contributed most to the analysis when they had strong correlations with the missingness and the amount of missing data was more than 25 %. In some cases, auxiliary variables were also helpful when the percentage of missing data was below 25 %. Overall, the study provided evidence that suggested that using more auxiliary variables, as opposed to restricting their number, was most helpful in reducing biases in the parameter estimates of the data model and improving efficiency in the parameter estimates. Taking a more liberal approach by including more auxiliary variables may help to ensure that particularly helpful auxiliary variables are included. They also noted that auxiliary variables need not correlate with the missingness but rather may be correlated with the data.

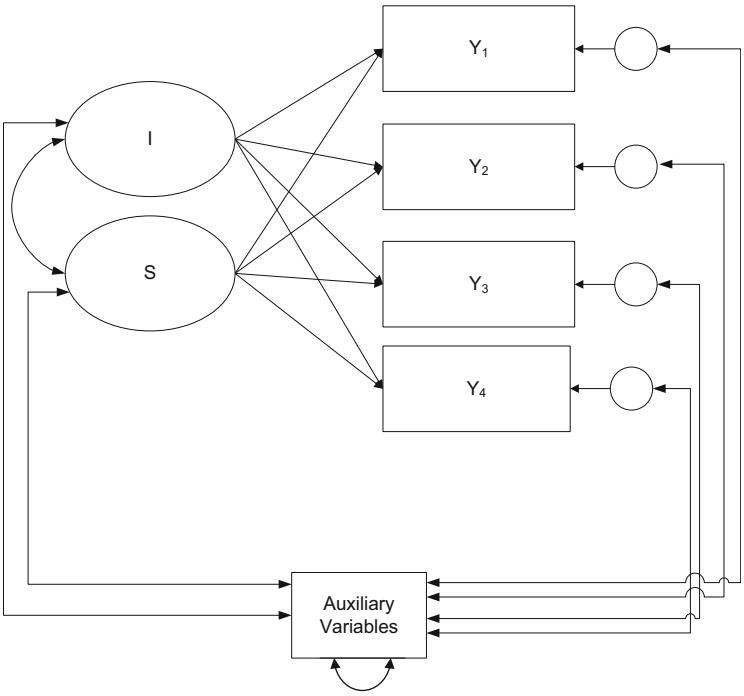

Fig. 24.4 Auxiliary variables in a saturated correlates model

Sensitivity of Parameter Estimates to Assumptions About Missingness

Unless data are generated with known characteristics such as by a simulation process, it is not possible to test for a given data set any assumption about the relationship between the missingness and the missing data. Indeed, any effort to evaluate any of the three data types (MCAR, MAR, and MNAR) can be carried out using only observed data. Thus, conclusions about the type of data involved in any case are only speculative. Given this, a recommended strategy in a data analysis when the missingness may not be ignorable is to perform what is called a sensitivity analysis.

Sensitivity analysis is often used to study whether changes in parameters estimates or statistical inference result after a model or the data have been modified, such as by making different distributional assumptions about data or making changes in the data (such as removing a case from a data set to study the influence of an individual's data on parameter estimates). In a data analysis involving missing data, a sensitivity analysis may be performed to evaluate whether parameter estimates depend on the specific assumptions made about the data (Little & Rubin, 2002). As noted earlier, a missing data process is not likely to be known

in practice. Thus, a recommendation is to consider multiple missing data treatments (Molenberghs & Kenward, 2007). This may involve applying different missing data frameworks (e.g., applying both a selection model and a pattern-mixture model), as well as formulating different missing data models (e.g., one model may assume that the missingness depends on the measured outcome at the occasion immediately prior to dropout, whereas another model might assume that the missingness depends on the measured outcome only at the start of the study).

Planning for Missing Data in Prevention Research

In prevention research, data analysis is very often done using methods that require complete data and so also require that data are MCAR. In other cases, methods that may make use of all of the observed data, such as structural equation models, require that data are MCAR or MAR. If whether or not data are missing is related to the missing values, however, inference from methods that require that data are MCAR or MAR can be problematic. If the missing data process is not ignorable, possible consequences include drawing false conclusions about the magnitude or strength of the relationship between measures or drawing wrongful conclusions about the effectiveness of a treatment program. Such problems may be avoided with careful planning in the beginning stages of data collection to address issues surrounding either potential sources of missing data or by making plans to measure variables, near or at the start of a study, that are likely to be correlated with the missingness or the missing data.

Summary

Prevention science research often results in data that are not complete. Many commonly used statistical procedures including regression analysis and analysis of variance generally require complete data. Consequently, missing data on at least one variable for an individual leads to an exclusion of the individual from analysis, resulting in a reduced overall sample size and statistical power, as well as increases in type I and type II errors. In some cases, analysis of only complete data can result in biased parameter estimates and wrongful statistical inference.

Given the need for multiple assessments over time in prevention science research, researchers are increasingly in need of appropriate statistical methods to handle missing data. Naturally, as the number of assessments over time increases, there is often a greater likeliness that some data will be missing, usually due to participant attrition. Structural equation models and mixed-effects models, methods that may rely on likelihood-based procedures for estimation, allow for missing data. These methods in their usual application require that data are MCAR or MAR. In a mixed-effects model, for instance, whether or not data are missing can depend on

the measured outcome when it is observed, such as observations made prior to the time when a participant drops out of a study. Conditional on this relationship, the missingness under MAR is assumed to be independent of the missing data (Laird, 1988). This is very different from a complete-data procedure, such as ANOVA, that under MCAR assumes that the missingness is independent of both the missing data and the observed data.

Structural equation models and mixed-effects models that address a nonignorable missing data process, as described here, have been well developed, and many of the procedures may be carried out using several commercially available software packages (e.g., Blozis et al., 2013; Xu & Blozis, 2011). A pattern-mixture random-effects model, for instance, may be used to examine differences in responses according to different patterns of missing data, such as testing whether the outcome measure for those who dropped from a treatment program versus those who completed the treatment changed on average according to different rates.

Multivariate analysis in general carried out using likelihood-based methods allows for the analysis of incomplete data so that all individuals may be retained for analysis. Likelihood-based methods, now a standard in the field, are valid when data are MCAR or MAR. In prevention science research, missing data can create an added challenge if data are MNAR. In these cases, analytic methods that assume data are MCAR or MAR may yield biased results. This paper reviewed state-of-the-art methods that may be used when data are MNAR. These methods include multiple imputation with auxiliary variables, selection models, shared-parameter models, pattern-mixture models, and models that include auxiliary variables that are correlates of the variables of a data model.

A major challenge in dealing with missing data is that assumptions about missingness cannot be completely tested. This is due to the fact that the missing data are not available for study. A preliminary analysis that suggests no differences between individuals with complete versus incomplete data with regard to background characteristics (e.g., gender, age) or the observed values of the measured outcomes, for instance, does not guarantee that data are MCAR because MCAR also specifies that the missing data are independent of the missingness. A common strategy for dealing with this problem of uncertainty concerning the status of the data is to perform a sensitivity analysis of the parameter estimates of a data model under different assumptions about the missing data.

A sensitivity analysis may involve the use of different missing data strategies, such as fitting different models that make explicit assumptions about the missing data process, as would be done when fitting a selection model, shared-parameter model, or a pattern-mixture model, or fitting a data model that has been extended to include correlates of the missingness or missing data. Whatever methods are selected for study, it is important to keep in mind that the study of missing data is carried out using only observed data. Thus, conclusions about whether the missing data process is ignorable or not are not certain. Furthermore, in cases in which the missingness is not ignorable but the true missing data process is not captured in the data analysis, it may not be possible to detect that a nonignorable missing data process is operating.

References

Blozis, S. A., Ge, X., Xu, S., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J., et al. (2013). Sensitivity analysis of multiple informant models when data are not missing at random. *Structural Equation Modeling, 20*, 283–298.

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*, 330–351.

Daniels, M. J., & Hogan, J. W. (2008). *Missing data in longitudinal studies: Strategies for Bayesian modeling and sensitivity analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Diggle, P. J., & Kenward, M. G. (1994). Informative dropout in longitudinal data analysis (with discussion). *Applied Statistics, 43*, 4–73.

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling, 10*, 80–100.

Graham, J. W., & Donaldson, S. I. (1993). Evaluating interventions with differential attrition: The importance of nonresponse mechanisms and use of followup data. *Journal of Applied Psychology, 78*, 119–128.

Graham, J. W., Hofer, S. M., Donaldson, S. I., MacKinnon, D. P., & Schafer, J. L. (1997). Analysis with missing data in prevention research. In K. Bryant, M. Windle, & S. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 325–366). Washington, DC: American Psychological Association.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11*, 323–343.

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods, 2*, 64–78.

Jöreskog, K. G., & Sörbom, D. (2006). *LISREL 8.80 for Windows (computer software)*. Lincolnwood, IL: Scientific Software International, Inc.

Laird, N. M. (1988). Missing data in longitudinal studies. *Statistics in Medicine, 7*(1–2), 305–315.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York: Wiley.

Molenberghs, G., & Kenward, M. G. (2007). *Missing data in clinical studies*. West Sussex, England: John Wiley & Sons, Ltd.

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles: Muthén & Muthén.

Rubin, D. B. (1978). Multiple imputations in sample surveys. In *Proceedings of the survey research methods section* (pp. 20–34). Alexandria, VA: American Statistical Association.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Rubin, D. B. (1997). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*, 473–489.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147–177.

Wu, M. C., & Carroll, R. J. (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics, 44*, 175–188.

Xu, S., & Blozis, S. A. (2011). Sensitivity analysis of a mixed model for incomplete longitudinal data. *Journal of Educational and Behavioral Statistics, 36*, 237–256. doi:10.3102/1076998610375836.

Part V
Cost Analysis and Policy

Chapter 25
Economic Analysis and Policy Studies: Special Challenges in the Prevention Sciences

Jonathan P. Caulkins, Nancy Nicosia, and Rosalie Liccardo Pacula

Introduction

There is growing pressure to view prevention interventions as "investments" in the public good and to hold them accountable for yielding favorable returns in the way that private investors expect financial investments to yield favorable financial returns. We call this enterprise "economic analysis" because the more familiar terms "cost-benefit" and "cost-effectiveness" refer to specific methods within this broader toolkit. And while "cost analysis" has also been used to name the broader toolkit, we prefer the label "economic analysis" to stress the importance of economizing the limited resources (e.g., money) available to society for these various programs.[1] Whatever the name, these expectations for accountability push beyond the familiar demands that programs be "evidence-based" and documented and maintain high fidelity with model program guidelines.

This chapter reviews core challenges that arise when applying such economic analysis to public policies and programs and are particularly salient when evaluating prevention interventions. We leave to others the tasks of both identifying which specific programs are most cost-effective (e.g., Aos, 2010; Aos, Leib, Mayfield, Miller, & Pennucci, 2004; Greenwood, Model, Rydell, & Chiesa, 1998) and raising the question of whether greater public investment in prevention generally is merited (e.g., Miller, Romano, & Spicer, 2000; Weissberg, Kumpfer, & Seligman, 2003).

[1] We do not use the term "economic analysis" in the sense that these methods belong to the discipline of economics, as this broad toolkit is not only or primarily used by economists. Moreover, only one of the tools in the toolkit (cost-benefit analysis) comes directly out of welfare theory in economics.

J.P. Caulkins (✉)
Heinz College, Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: caulkins@andrew.cmu.edu

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_25,

We write for two audiences: those who know a lot about prevention but not so much about economic analysis and experts in economic analysis who lack substantive expertise about prevention science. For the former, we explain how these alien, almost mystical methods work so well; for the latter, we explain why these powerful, almost majestic methods do not always work so well.

To avoid any misunderstanding, we will state our position explicitly at the outset. We believe prevention interventions should be subject to the scrutiny of economic analysis. However, analysis that seems straightforward in concept can become messy in its particulars. We also believe that dogmatic or formulaic application of economic analysis that does not make common sense adjustments for the practical realities of prevention interventions can mislead rather than enlighten.

The essential premise of economic analysis in the service of public policy is simple: Social welfare is maximized through the efficient allocation of resources. In theory, this maximization occurs simultaneously across all choices a society faces, and efficient allocation means that no person can be made better off without making someone else worse off; this is known as the "Kaldor-Hicks Criterion" (Boardman, Greenberg, Vining, & Weimer, 2006). Cost-benefit analysis is the particular tool solidly rooted in that theory. We will use the adjective "allocative" when we are referring to efficiency in the Kaldor-Hicks sense. In practice many if not most decisions are informed by less ambitious notions of efficiency, for which a range of additional tools become relevant.

In either case, there is general agreement that efficiency analysis should consider *all consequences* for society that are *caused by* a decision or program within a unified analysis. When considering a single program, the recommendation is a "go" if and only if the *present value* of cumulated effects over time, net of *all costs including opportunity costs*, is positive. When there are multiple programs, they should be compared on some *common metric*.

Of course, efficiency is not all that matters when making decisions with public dollars. Governments appropriately act for other reasons including justice, national security, and compassion; if a child is trapped at the bottom of a well, we do not calculate the cost of rescue before deciding whether to act. Yet efficiency is widely accepted as an important desideratum and economic analysis is just a collection of procedures that seeks to support that objective. It is common to say that in these times of budgetary challenges we must pay particular attention to cost-effectiveness, but it is the exception when a society is so affluent as to not care about efficiency.

Nevertheless, economic analysis remains controversial. It has variously been decried as inhuman, irrational, and/or inane. The resistance has two roots: (1) lay misunderstanding of sensible, albeit esoteric, procedures and (2) hubris among its methodological experts concerning fundamental limitations of its procedures. The five emphasized phrases in the paragraph above account for the bulk of what makes economic analysis provocative for prevention science, in both good and bad ways. So, after an introductory vignette, we organize this discussion around those five themes in the following order: "common metric", "all costs including opportunity costs", "present value", "caused by", and finally "all consequences".

Introductory Vignette

It is helpful to connect economic analysis for public policy with that in the business world. Suppose a company had the following option. It could spend $10 million on a marketing campaign that would induce 100,000 people to buy its product. Should the firm proceed?

In order to answer that question, it is necessary to quantify the value of an incremental sale. Suppose the wizards in the accounting department determined that the firm's profit margin on each sale is $200. Then economic analysis would say "yes," launch the marketing campaign. At $200 per customer, generating 100,000 more customers is worth $200 × 100,000 = $20 million, and it is sensible to buy something worth $20 million if it only costs $10 million.

Similarly, suppose a public health agency could spend $10 million on a vaccination campaign that would prevent 100,000 cases of flu. Should the agency proceed? It is hard to say because it is not obvious how to value preventing one case of flu. Cases of flu are neither traded on a commodities exchange nor bought at the local mall, so how does one attach a price or value to there being one fewer case of flu?

Economists have grappled with such questions for decades and developed a variety of creative answers. We'll elaborate below, but foreshadow here. One can, of course, ask people about their willingness to pay to avoid an outcome, known as "contingent valuation." If the average person says he would pay up to $200 to avoid the flu, then perhaps $200 is a good estimate of the value to society of preventing a case of flu. But people tend not to be very good at answering hypothetical questions, particularly about things less familiar than the flu. When given the chance to actually spend their own money to avoid an adverse outcome, people often spend much more or much less than what they self-report as their willingness to pay. So economists prefer to base valuations not on mere words, but instead on actions observed from real market transactions or "revealed preference." For example, the amount people spend on burglar alarms provides insight into how much they dislike being burglarized.

Suppose that the wizards in the economics department determine that the relevant population's revealed preferences indicate that it values not getting the flu at $200 per case of flu. If we ignore for now complications such as the possibility of side effects, herd immunity, and the like, then the analysis is the same as that for the for-profit firm and so is the conclusion. The public agency should be willing to spend $10 million in order to prevent the public from suffering $20 million worth of flu.

Thus, economic analysis views averting a cost and generating a gain as equivalent, which is good, because almost by definition most of the benefits of *prevention* programs come from preventing costs, not generating gains. So in this theoretical sense, economic analysis does not discriminate against prevention programs. However, as a practical matter, it is much harder to count events that do not happen than it is to count events that did happen. We return to the problem of estimating the causal consequences of prevention below.

The examples above were either/or decisions: either launch the marketing/public education campaign or not. In the jargon, those are "go/no-go decisions." Usually, however, there are more alternatives. For example, suppose the firm or public health agency could consider an alternative initiative, one that costs $25 million but whose reach was 200,000—either new customers for the firm or flu cases averted for the public health agency. The net benefit of this new program is greater because 200,000 × $200 − $25 million = $15 million—which is larger than the $10 million net benefit from the first program. So, given a choice, both the firm and the public health agency should opt for this second initiative because it offers the larger net benefit.

This vignette also illustrates the difference between thinking in terms of net benefits versus a cost-effectiveness ratio. Someone might object that the original program was more cost-effective. It reached 10,000 people per $1 million, whereas the alternative program's corresponding ratio is only 200,000/$25 million = 8,000 people per $1 million. That is true, but irrelevant if the programs as described are the only options available. However, if the original program could be scaled up proportionally, so that by spending $25 million (2.5 times as much as originally proposed), one could reach 250,000 people (2.5 times as many), then it would be preferred because its net benefit would then be 250,000 × $200 − $25 million = $25 million, which is even better than the $15 million net benefit of the alternative program.

Let's push this example one step farther to highlight an important difference in perspective between the firm and the public health agency. When a firm spends $10 million to make $20 million, it ends up with more cash in the bank than it started with. It is a single entity that makes the decision, pays the cost, and receives the benefit. Moreover, all decision makers within the firm have the same objective, to maximize profits, and hence decisions are more easily made.

Things are more complicated for the public health agency. The agency spends from its own budget to create nonmonetary benefits that accrue to others. So, unlike the firm, after running a highly cost-effective campaign, it has less cash in the bank (or, more realistically, less spending authority left under its budget). It would be nice if the people who didn't get the flu expressed their gratitude by writing the agency a $200 check, but they don't. In reality, these people likely don't even know they were among the 200,000 beneficiaries who did not get the flu due to the program, let alone know who they should thank for that valuable service. They might mistakenly credit good genes, their personal trainer, or vitamin supplements.

In theory, for countries with governments of the people, by the people, and for the people, we are all one big happy family or "society." We the society decide, we the society pay, and we the society benefit from the flu prevention campaign, so from society's perspective, the analysis is equivalent to that for the firm. But, in practice, the world is full of bureaucratic silos and parochial perspectives. Benevolent social planners seeking to maximize social welfare are few and far between; consider that in the entire United States, there are only two government employees who are even charged with looking after the country's overall welfare (the President and Vice President).

So economic analysis of public policy needs to be thoughtful about from whose perspective the analysis is done. "Society as a whole" is the most common perspective adopted, particularly in academic work, but it is not always the most useful or influential to the extent that it is an abstract ideal, divorced from more tangible priorities of actual decision makers (Elvik, 2010; Humphreys, Wagner, & Gage, 2008). Good researchers often address the problem by evaluating the same program from multiple perspectives (e.g., Lee, Aos, & Miller, 2008).

The (Elusive) Common Metric

The vignette introduced two broad categories of cost analysis metrics: (1) net measures, which capture the difference between benefits and costs, and (2) ratios of two measures of interest such as cost per benefit (or benefit per cost). The former use subtraction and the latter division.

The primary advantage of using the difference between benefits and costs (commonly referred to as "net benefit") is that it makes it simple to apply the Kaldor-Hicks criterion of implementing only those programs for which a net gain is obtained by society. The main drawback of focusing on net benefits is that both terms have to be measured in the same units; as your grade school teacher explained, you are not allowed to add (or subtract) apples and oranges.

The principal benefit of using ratios is that there is nothing wrong with taking ratios of things measured in different units. Every driver understands the meaning of miles per hour and miles per gallon, or their metric equivalent, even though it makes no sense to add two miles plus two gallons. That flexibility is valuable when it is practically or politically difficult to translate one unit into another. If spending $10 million prevented 100 rapes instead of 100,000 cases of flu, then computing the program's net benefit would require defining the value of preventing a rape. Mathematical continuity implies there must be such a number somewhere between $0 and a million times the GDP of the entire world, but that does not mean any government official wants to endorse an analysis that names such a number.

The more common unquantifiable is life. How much is a human life worth? Volumes have been written on the topic, and some experts will argue long and hard that this or that approach is the "right" way to value a human life in cost analyses. Nevertheless, whenever possible the prudent analyst and every politician will compare two life-saving programs in terms of their cost-effectiveness (e.g., lives or life years saved per $1 million) not on net benefit terms.

The catch is "whenever possible." Suppose one must choose between two crime prevention programs. Each costs $10 million. One will prevent two homicides and ten rapes; the other will prevent four homicides and two rapes. The first is more cost-effective at preventing rape; the second is more cost-effective at preventing homicides. The ideals of economic analysis compel us to consider all important consequences, and clearly rapes and homicides are both important. So sometimes

there is no getting around the need to value disparate outcomes in some common metric.

With that introduction, let us now define the primary performance metrics used in policy analysis by briefly describing the four basic techniques: cost offsets, cost-effectiveness analysis (CEA), cost-utility analysis (CUA), and cost-benefit analysis (CBA). While all of the techniques are used regularly in policy evaluation, and termed as "economic evaluations," only CBA adheres to the principles of welfare economics and the goal of allocative efficiency, which is why it is the preferred method of economists.

Cost Offsets

Cost offset analysis counts not only costs generated by the program[2] but also cost savings (or revenue gained) accrued to the agency paying for the program. A classic example would be providing preventive care that obviates the need for more expensive treatment (e.g., screening for hepatitis, so treatment can begin before the disease has progressed). Sometimes the cost offsets (reductions) are bigger than the original program costs, so implementing the program actually saves money. An example of such a program in the drug prevention area is raising the tax on beer to reduce alcohol initiation and/or binge drinking. The cost of imposing or raising the tax on alcohol is more than fully offset by the revenue generated from the higher taxes.

Cost-Effectiveness Analysis

CEA compares ratios of the cost of a program relative to its effectiveness at producing a specific desirable outcome (Drummond, O'Brien, Stoddart, & Torrance, 2005; Gold, Siegel, Russell, & Weinstein, 1996). For example, if the key outcome were overdoses averted, then

$$\text{Cost Effectiveness Ratio} = \frac{\text{Cost w Intervention} - \text{Cost w/o Intervention}}{\text{ODs averted w Intervention} - \text{ODs averted w/o Intervention}}$$

Thus, the CEA ratio indicates how much one has to "pay" to "purchase" one unit of the benefit (e.g., cost per overdose averted). Expressed in this format, smaller ratios are better, indicating less expensive ways of purchasing a given benefit. Some

[2] We will henceforth use the term "program" to refer to a program, policy, or intervention that might be considered by the government or prevention service provider.

literatures prefer the inverse ratio so that larger numbers are better (more benefits purchased per million dollars spent).

CEA measures program desirability in a way that can be compared directly across competing programs but is limited to interventions that produce one primary benefit. In some instances, decision makers might only be concerned about the gains in terms of one primary measure of interest (lives saved, students graduating), but for many areas of policy, the benefits (or cost offsets) cannot be easily represented along a single dimension. Substance abuse prevention programs represent a perfect example in that many prevention programs have components that also improve self-esteem and self-advocacy and reduce peer influence thereby reducing delinquency and incidences of unprotected sex and improving rates of graduation (Ellickson, Collins, Bogart, Klein, & Taylor, 2005; Hawkins, Kosterman, Catalano, Hill, & Abbott, 2005; Mensch & Kandel, 1988).

Cost-Utility Analysis

CUA combines multiple benefits into a single aggregate measure when all of those benefits are health related (Drummond et al., 2005). Like CEA, CUA considers the costs and outcomes simultaneously by forming and comparing ratios. The outcome of interest, however, is expressed in expected utility or general well-being with respect to health, not a single outcome such as averted overdose. Changes in utility are expressed in healthy-year equivalents (HYEs), disability-adjusted life years (DALYs), or, most commonly, quality-adjusted life years (QALYs) (Gold et al., 1996). These metrics all assign a numerical weight to each health state that reflects its desirability, ranging from zero (death) to one (optimal health). Intervention effects are represented as changes in health states, and the quality-adjusted score for each health state is multiplied by the expected time spent in the state and then summed over the expected time of life.

Eliciting preference weightings for health states is difficult. Specialists in this area grapple with issues like the fact that prospectively people think an adverse health outcome (e.g., loss of a limb) will be much worse than do people who have suffered that outcome and have had time to adjust to it. Two widely adopted methods are the healthy days measures (Centers for Disease Control & Prevention, 2000) and the EuroQol-5D (EQ-5D) (Kind, Dolan, Gudex, & Williams, 1998; Rabin & de Charro, 2001). Neither is perfect in its measurement of health states. Furthermore, neither reflects the non-health aspects attributable to prevention, such as reductions in criminal engagement or child neglect associated with drug addiction. Hence, scientists are still working to improve measures of well-being such as the "quality of life index" (QOLI) (Daley, Shepard, & Bury-Maynard, 2005; French, Salome, Sindelar, & McClellan, 2002). Fortunately, as improved measures emerge, they can easily be folded into the same basic analytic framework.

Cost-Benefit Analysis

Prevention programs often improve a range of non-health outcomes in addition to improving health: Drug prevention programs might lead to better family relations, improved employment, and less criminal involvement. Other policy domains also recognize that health is not the only relevant criteria. For example, policy analysts evaluating a proposed dam would factor in outcomes such as electricity production, enhanced recreational opportunities, and flood control (including protection of property as well as lives). Hence, other than in the domain of health, policy analysts generally favor CBA as they frequently are forced to consider multiple, distinct outcomes. Like CUA, CBA attempts to convert all outcomes into a common metric, but in this case, the common metric is currency (dollars in the United States, euros in the Eurozone, etc.) rather than the desirability of health states.

Just as with CUA, there is a literature devoted to determining how to assign reasonable values to various outcomes. For example, what is the societal cost of a burglar stealing something? The old answer was zero on the grounds that theft merely transfers wealth from one person to another, with no net loss to society. However, the public and policy makers accountable to the public vehemently object to the idea that there is no social cost to property crime. The more modern approach is to focus on the "willingness to pay" (WTP) to avoid such criminal victimization. That is, if an average person were willing to pay $1,000 to avoid being burglarized, then preventing a burglary should be valued at $1,000.

Likewise, for violent crimes, the older approach focused on victims' hospital bills and days of work lost while recuperating, but this led to valuations that struck the public as absurdly low. Estimates based on jury awards and WTP principles that incorporate the "intangible" costs of pain and suffering are much higher.[3]

One might think crime costs are important for law enforcement interventions but a digression for prevention. However, crime is so costly to society that its prevention can account for an important part of the justification for interventions that have nothing directly to do with policing. For example, they account for half of Lee et al.'s (2008) estimated benefits of nurse home visits for low-income families expecting their first child. In the area of substance abuse prevention, crime is consistently the second largest component of the social cost of alcohol and drug abuse (Harwood, Fountain, & Livermore, 1998; Office of National Drug Control Policy, 2004).

The classic imponderable is how does one value a human life? The contentious human capital approach values human life based on the present value of expected future earnings (Rice & Cooper, 1967), so it values the lives of those with high-paying jobs more than those of individuals who choose not to work (e.g., stay-at-home

[3] Cohen (2005) and Miller, Cohen, and Wiersema (1996) are standard references for the valuation of crimes prevented.

parents) or who earn lower wages. The WTP approach gets around this issue by calculating the value of a statistical life (VSL) based on purchases people make to reduce their risk of death, or premiums they must pay to accept additional risk of mortality, and applying the same rate to all people in that society (Schelling, 1968; Viscusi & Aldy, 2003). Examples of such transactions include the purchase of safety devices (e.g., cars with air bags, home security systems) and wage premiums for working in risky jobs (e.g., mining or construction).

Before sighing in relief at this apparently successful evasion of distasteful inequity across income groups, note that applying the same value to all lives lost automatically assigns a much lower value to a year in the life of someone who would have died young (e.g., prevented drug overdose) than in the life of someone who would have died when older (e.g., prevented death from seasonal flu). Conversely, applying the same value to all life years lost automatically assigns a lower value to deaths of older people than of younger people. Considerable attention has been devoted to the impact of age effects in analyses of VSL by economists, with results showing how VSL and age depend on how well capital markets function (Aldy & Viscusi, 2008; Johansson, 2002; Shepard & Zeckhauser, 1984).

Don't let jargon and acronyms lull one into thinking that economic analysis is value-free. In reality, seemingly minor technical choices by the analyst, such as the choice to value outcomes in terms of life years rather than lives lived, do involve valuation judgments that indeed can affect which program comes out the winner in an economic analysis.

A benefit of particular relevance to prevention sciences whose quantification is in nascent stages is prevention of drug dependence.[4] There are health problems correlated with dependence for which estimates have been made, including hepatitis, depression, and living with AIDS. But imagine a person who is in perfect health except for being dependent on cocaine. QALY losses are routinely considered for other psychiatric disorders. Dependence is a recognized medical condition, and most people would view drug dependence as a state with a lower quality of life than being in perfect health and not being dependent; the cravings, obsessing about obtaining the next dose, and general loss of control adversely affect daily functioning, just as surely as would other ailments. But the social cost of dependence per se is rarely included in economic analyses.

Rough calculations suggest this omission could be problematic,[5] but many economic analyses still assume that even dependent users are like conventional

[4] Another is the psychological pain and suffering of family and friends of those who are addicted.

[5] Some authors assign a loss of 0.1 QALY per life year spent drug dependent (Stouthard et al., 1997), and that seems consistent with one of the few empirical studies (Pyne et al., 2008). ONDCP (2004) estimates there were 2.7 million "hardcore" cocaine users in the United States in 2000, suggesting a loss of 0.27 million QALY per year from cocaine dependence. A traditional but dated valuation is $65,000 per QALY (e.g., French et al., 2002), but updating with more modern estimates of the value of a life (e.g., Viscusi & Aldy's, 2003 valuation of $4–$9 million per life) suggests valuations closer to $250,000 per QALY. So the cost of cocaine dependence itself, as distinct from its sequelae, could be on the order of $17–$70 billion per year, which is not small compared to the cost-of-illness study estimates of drug-related costs (e.g., Harwood, 2000).

consumers in the sense that the satisfaction derived from consuming a product (drugs in this case) can be assumed to exceed its purchase price; otherwise, the person would not have consumed it. By that logic, not only could there be no cost of dependence, one should actually count drug use as a benefit, so preventing use creates a negative benefit (for our purposes, a cost).

Intoxication, as distinct from dependence, presents a similar challenge. The economic costs of alcohol use are large; one frequently referenced figure is $185 billion in the United States in 1998 (Harwood, 2000). Americans now consume on the order of 100 billion standard drinks per year, meaning 0.6 ounce servings of alcohol whether in the form of 12 ounces of beer, 5 ounces of wine, or 1.5 ounces of spirits (Cook, 2007). Someone morally opposed to intoxication might assign a social cost of a few dollars per drink, while an economist committed to valuing material pleasures might assign a benefit of a few dollars per drink. If either valuation of alcohol consumption and intoxication (the negative or the positive) were used in calculating the efficiency of alcohol prevention interventions, it would loom quite large in the analysis.

Likewise, Kilmer, Caulkins, Pacula, MacCoun, and Reuter (2010) estimated various budgetary consequences of California's narrowly defeated proposition to legalize marijuana. Many outcomes that attracted considerable attention in the debate had magnitudes of $5–$200 million per year. Yet Kilmer et al.'s (2010) figures suggest current consumption in California is roughly one billion joints per year, a level they judged could plausibly have doubled with legalization. So for someone assigning a benefit or a cost to the increased marijuana consumption, whose absolute value might be $1 or more per joint consumed, the monetized value of increased intoxication dwarfs the more commonly discussed outcomes.

Resolution of these issues awaits further research, but the discussion illustrates two important points: (1) reductions in use and dependence may be an important contributor to the benefits of prevention programs, in at least some observers' eyes, if they are not dismissed summarily on theoretical grounds and (2) quantification of benefits may inevitably be more nebulous and more controversial for prevention programs than for firms, particularly in domains for which the public believes that the activity itself, not just its consequences, is a matter of public concern.

All Costs Including Opportunity Costs

The previous section makes clear that monetizing benefits can be tricky. One might think that monetizing program costs would be a breeze. That may be true by comparison, but only by comparison. A surprising number of complications arise just in quantifying costs. Many stem from the intersection of two facts: (1) prevention evaluations usually take a societal not an agency perspective and (2) many of the resources employed are donated in kind, not purchased.

Analysis of private-sector investments is simpler. For the most part, business analysts only care about outcomes that affect their company, so they do not worry

about tracing out indirect or spillover effects. And, few people make charitable donations to for-profit companies. So almost all resources consumed come with a "purchase receipt" (payroll records in the case of "purchasing" employees' time). We do not mean to belittle the hard work of MBAs; they do face complications, such as how rapidly one should depreciate a capital asset. But it is a lot easier to figure out what something is worth if it is bought and sold in arm's length transactions.

Contrast that with, say, a mentoring program that matches adult volunteers with at-risk youth. The most valuable resource consumed is the donated volunteers' time. The conventional approach assigns a shadow or opportunity cost to those hours, such as the wage of professionals providing similar services or the wage of the volunteers themselves in their primary employment. The argument for doing it this way is that societal resources are inherently limited, and when they are used for one purpose, they can no longer be used for another. So the wage the volunteer could have earned doing a similar job represents the opportunity cost of his or her time and is the true economic value of the labor. But those wages could be very different when a corporate lawyer volunteers to read to children in a daycare program than when a social worker volunteers.

One could argue instead that volunteers must be deriving positive net utility from their participation; otherwise they would not have volunteered. Indeed, one could go a step farther and conjure up arguments that society is better off when citizens become involved in one another's lives, so the volunteer hours not only benefit the volunteers and children but also create a broader set of positive externalities. How to reconcile such divergent views is not something even considered by Fortune 500 companies' investment strategies because they are irrelevant, but specific views on how to value these resources could directly influence whether mentoring programs are viewed as among the most or the least cost-effective forms of prevention.

Similarly, many media-based prevention campaigns employ donated air time, school districts donate facilities to after-school programs, and they donate not just classroom space but also classroom time to school-based programs. (Classroom space is donated in the sense that most prevention programs do not write checks to the school district to "rent" the classroom the way a corporation rents a conference facility to host a corporate training event.)

To underscore how complicated this can be, consider the Drug Abuse Resistance Education (D.A.R.E.) program, whose curriculum is delivered in schools by police officers. The schools donate space, the police donate the instructor's time, and the nonprofit organization D.A.R.E. America donates the curriculum materials—at least for the first year. Donations raise few challenges for analysis undertaken from the perspective of a single agency (e.g., a public school district); for the agency, costs are mostly reflected in net cash flow and donations do not show up in the flow of dollars.

Usually, however, prevention interventions are evaluated from a societal perspective, and society values the donated resources even if the agency administering the prevention program does not have to purchase them explicitly. When the

National Highway Traffic Safety Administration (NHTSA) does a CBA of installing air bags in cars, it counts the cost of the air bags (reflected in the higher price of new cars) not only for the few cars the agency purchases for its own purposes but for all new vehicles purchased in the country. In effect, car buyers are donating the money that funds the improvement in vehicle safety, and NHTSA rightly recognizes that as a cost to society, even though it never shows up in the agency's budget.

The same line of reasoning applies to a range of unfunded prevention mandates, including motorcycle helmet laws, zoning ordinances that reduce the risk of fires, and many Occupational Safety and Health Administration rules.[6]

So, from a societal perspective, it doesn't matter who bears the costs or donates the inputs because everything is accounted for. Standards are emerging governing how to value various tangibles, such as facility space, but the elephant in the room is valuing the time of the prevention intervention's target audience. Caulkins, Everingham, Rydell, Chiesa, and Bushway (1999) find that the opportunity cost of giving youth a school-based prevention program, instead of using that classroom time to teach conventional academic subjects, dwarfs the budgetary costs for program materials and training that are the traditional focus of cost estimates (e.g., Drug Strategies, 1996). They suggested valuing those classroom hours at the average cost of providing secondary education but noted that could be conservative inasmuch as most voters believe public education is a good use of taxpayers' dollars, meaning the benefits of education exceed the cost. Others might argue that students' time has no value because they are too young to earn a wage in the labor market or that the teacher's salary is a better surrogate. We do not want to delve into the pros and cons of each of these approaches. Rather, our point is that thoughtful people can disagree, and the consequences of their choices can matter enormously.[7]

An interesting corollary is that sometimes the traditional objects of cost data collection can be ignored, if it can be shown a priori that they will be swamped by other factors. There is no sense estimating the cost of program materials to the penny if such precision will be swamped by uncertainty about how to value the opportunity cost of using class time for a prevention intervention.

Although generally not discussed, one could raise similar questions about other prevention modalities such as "midnight basketball leagues" (more formally,

[6] We spend enough time in airports to wonder if the cost to bystanders of exposure to incessant inane warnings about the approaching end of the moving sidewalk generates stress-related mental health problems that more than offset any benefits from prevented moving sidewalk falls.

[7] An interesting corollary is that a comprehensive evaluation of school-based programs ought to assess effects on academic outcomes. Consider two programs that involve the same number of classroom contact hours and the same (minor) out-of-pocket costs. The first has zero academic value but reduces drug use by some amount. The second reduces drug use by only three fourths as much but produces half as much general education benefit as the academic classes displaced by the prevention program. The second program would lose if the analysis considered only the effects on drug use and those on general education were ignored, but it would win in a more comprehensive economic analysis.

sports-based community prevention). Even leaving aside the considerable challenges of estimating benefits (cf. Hartmann & Depro, 2006), what are the costs? If the courts already exist and would otherwise be idle (so no meaningful opportunity cost for facilities) and the adults are not paid staff but volunteers who enjoy shooting hoops and spending time with kids, decisions about valuing people's time could well dwarf the tangible budgetary costs for some posters and a few basketballs.

If government oversight agencies allow methodological latitude in how analyses value donated time and other resources, the door is wide open to the nonprofit world's equivalent of corrupt creative accounting practices. But if oversight agencies tightly prescribe the accounting rules, they risk distorting incentives. For example, if an agency were held firmly to the principle that the opportunity cost of volunteers' hours must be valued at their outside wage, that would create perverse incentives to reject highly educated volunteers because they are more costly. For example, a midnight basketball league could suddenly appear to be extremely costly if an NBA star decided to volunteer.

Present Value

One of the most common acronyms in economic analysis is NPV (net present value). Policy analysts of all stripes agree that if the costs and outcomes do not occur at the same point in time, they must be "discounted" back to their "present value." Discounting—or the failure to do so—can be particularly consequential when evaluating prevention interventions because there can be very long lags between program spending and program outcomes. Running a teen smoking prevention program today may prevent a cancer death 50 years down the road.

With money, it is easy to see that a dollar or euro today is not the same as a dollar or euro tomorrow, let alone in 50 years. Normally a given amount of money today is worth more than the same amount in a year because if we had the money today, we could invest it and have more money in 12 months.

In particular, it is common to assume that a dollar today is worth $(1 + r)$ dollars next year and $(1 + r)^n$ dollars in n years, where r is the annual rate of return on available investments. Conversely, a dollar in n years is worth only $(1 + r)^{-n}$ dollars today. Because the rate of return is (essentially) always positive, $(1 + r)^{-n}$ is always less than 1, so money in the future is worth less than the same amount of money today. Hence, to translate dollars in the future into dollars today, we "discount" the future dollars back to their "present value."

So far this sounds interesting only for those who sport green eye shades, a mere technocratic adjustment. What makes discounting controversial is that it is, and in fact should be, applied to nondollar outcomes, including such sacred things as lives saved. This is a strange concept whose name creates a public relations nightmare. If one asks, "Should we discount future lives?" anyone with a whiff of common sense would say, "Of course not." But he would be wrong. Keeler and Cretin (1983) prove

that we must discount nondollar outcomes and at exactly the same rate; failing to do so can lead to bad decisions, wasted funds, and lives not being saved.

The gist of their argument is captured in the following thought experiment. Suppose we have the option of funding a wonderful prevention intervention; for every $100,000 we invest today, we will save one life in 10 years. (To keep things simple, suppose the program is politically popular, generates no adverse indirect effects, etc.) That opportunity—call it Option A—would generally be thought of as a terrific bargain, and any reasonably affluent society would wish to fund it.

Now most programs that can be run today could just as well be run the following year. So presumably we also have the option—call it Option B—of spending $100,000 in 12 months and thereby saving one life 11 years from now. One way to implement Option B would be to put $97,000 in the bank, let it earn 3 % interest over the next year, and then withdraw it to pay for the program.

If we do not "discount" the future lives saved, the choice is easy. Option A lets us save one life for $100,000, but Option B saves the same number of lives for only $97,000. So clearly we should choose Option B; it costs less. Or, assuming these interventions are scalable, we might spend $100 million on Option A today and save 1,000 lives, or we could put $100 million in the bank for a year, use the proceeds to invest $103 million in Option B next year, and thereby save 1,030 lives. Either way, Option B wins.

Of course if we put the $100 million in the bank, then 12 months later we will realize there is an Option C. Namely, leave the money in the bank for another 12 months, letting it accrue further interest until we spend it 24 months later to save even more lives. And so on, ad infinitum.

So we have a paradox. No matter how cost-effective a prevention program is, if we do not discount future nonmonetary benefits, we will always do nothing, preferring to put the money in the bank, saving for a future investment that it is never optimal to make.

Surely if we are impatient and prefer to have our money today, we ought to feel the same urgency about preferring to save lives now, rather than dawdling, but unfortunately this natural impulse to want to implement life-saving measures sooner rather than later gets translated in the emotionless jargon of economics into the cold-hearted phrase, "discounting the value of future lives." Perhaps no phrase does more to discredit the entire enterprise of economic analysis, but its roots are entirely humanistic. It is merely the logical consequence of insisting that we be as inpatient in our desire to save lives as we are in our desire to receive money.

Even after one gets comfortable with the idea of discounting future outcomes, there remains the question: What discount rate should be used? When a company assesses investment opportunities with discounted cash flow analysis, it is generally accepted that future outcomes should be discounted at its WACC or weighted average cost of capital. That is essentially its cost of borrowing money. If the business would raise cash to make an investment by taking out a bank loan at 8 % interest, then it ought to use an 8 % discount rate in its analysis. In a typical situation of an upfront investment generating positive cash streams in the future, this makes

the usual go/no-go decision rule ("proceed if discounted benefits exceed discounted costs") equivalent to "invest in your own operations if the internal rate of return exceeds the interest rate on the loan that could be used to finance those operations," an eminently sensible approach.

But public agencies do not usually fund prevention programs by taking out bank loans, so what discount rate should they use? This crucial question was debated vigorously in the 1960s and 1970s, when economic analysis was first commonly applied to public investments. The resulting literature is enormous, provocative, carefully reasoned, and ultimately inconclusive. For example, there have been heated debates about whether it matters if the benefits accrue to a different generation of people than those paying the bills, whether one should pay attention to the likelihood that future generations will be more affluent than we are today, and whether one should inflate discount rates when outcomes are uncertain. Had you asked three smart and well-trained economists in 1980 what discount rate to use you might well have received four answers.

The tendency over the last few decades has been for literatures to just pick some value and stick with it—the benefit being that it improves comparability across analyses. Indeed, one recent contribution to this literature is delightfully entitled "Just Give Me a Number!" (Moore, Boardman, Vining, Weiner, & Greenburg, 2004). For example, the Office of Management and Budget advocates using a "real" rate of 7 %, and the classic reference for economic analysis of healthcare interventions (Gold et al., 1996) suggests doing the analysis twice, once with a 3 % and once with a 5 % discount rate to see whether the results are robust with respect to this parametric choice.[8]

This détente is not fully satisfactory for two reasons. First, while in many domains it does not usually matter whether one uses a 3 %, 5 %, or 7 % discount rate, for prevention program whose gains are far off in the future while the costs are immediate, simple arithmetic shows that those seemingly small distinctions can have very large impacts.

Suppose, for the sake of argument, that we are willing to fund any prevention program whose cost-effectiveness, in present value terms, is at least one life saved per $2 million spent. And, returning to the example with which we opened the section, suppose that spending $100 million on teen smoking prevention programs today would prevent 1,000 cancer deaths 50 years down the road. Discounting at 3 % versus 7 % decisively alters the program's cost-effectiveness. With a 3 % rate, the cost per (discounted) life saved is only $400,000, readily meeting the efficiency test, whereas with a 7 % discount rate, the cost rises to $3 million, and the program would not be approved.

[8] Fortunately concerning the odd-sounding distinction between "real" and "nominal" discount rates, naïve impulse aligns with logic; we should focus on the "real" rates, meaning the rate after adjusting for inflation. For example, if the actual or "nominal" interest rate is 10 %, but inflation is running at 3 % per year, then the "real" interest rate is only $10\ \% - 3\ \% = 7\ \%$.

When benefits accrue within a shorter time frame such as 5 years, the choice of discount rate matters, but not enormously. Moving down (up) from a 5 % to a 3 % (7 %) rate increases (decreases) the cost-effectiveness ratio by a factor of 1.1. But over the longer time horizons that arise routinely for some prevention interventions, discounting generally and the choice of discount rate in particular become tremendously consequential choices.

Second, there is now clear evidence that individuals do not apply a constant discount rate over time as in standard "exponential" discounting. They might discount in another way, such as by "hyperbolic" discounting (Laibson, 1997).[9] There is evidence that societies do as well (Henderson & Langford, 1998). Some suggest this is normative, not just descriptive, and calls for hyperbolic discounting to be done in economic evaluations because the social discount rate should in some sense reflect an aggregation of individuals' (heterogeneous) time preferences (e.g., Nocetti, Jouini, & Napp, 2008). This has become an active topic of discussion in part because traditional exponential discounting discourages taking action to prevent global warming (Karp, 2005). Drug and violence prevention advocates should cheer on environmentalists in this regard, but for the time being, the state of the art is applying one somewhat arbitrary but conventional per-period discount rate to all time periods.

Caused By

Perhaps the least controversial bedrock principle of economic analysis is that only those consequences that are caused by a prevention program should be stacked up against its costs. In many disciplines establishing causal connections is not controversial. The rules of Newtonian physics are so powerfully predictive that no one argues about what causes billiard balls to move. In others—including much of clinical medicine—we believe in causality if and only if it has been demonstrated via a randomized clinical trial (RCT).

RCTs in social science are becoming increasingly common, in light of the early successful trials in healthcare insurance (Newhouse, 1993), poverty (U.S. Department of Health and Human Services, Administration for Children & Families, 2010; U.S. Department of Housing and Urban Development, Office of Policy Development & Research, 2003), and criminal justice (see Farrington & Welsh, 2005, for a review). RCT studies play an important role for many prevention interventions, but they are not practical in others. Consider the very important domain of preventing war, which includes an idea called "democratic peace

[9] What exactly hyperbolic discounting means is not important here, but for those not familiar with the term it essentially means dividing by $(1 + rn)$ rather than $(1 + r)^n$, albeit for a different value of r. The key point is that for outcomes in the medium to distant future, the distinction between hyperbolic and exponential discounting is highly consequential.

theory." Roughly speaking, the hypothesis is that democracies rarely if ever go to war with each other, from which some might extrapolate the idea that spreading democracy causes peace.

Is it reasonable to expect proponents of democratic peace theory to bolster their beliefs with evidence from an RCT? It is not clear whether institutional review boards would sanction randomly assigning the people in some countries to receive democratic representation and others to live under dictatorship, or who would have the authority to implement the randomization, let alone how to keep the whole experiment double blind.

This example is, of course, extreme, but there are environmental prevention interventions for which random assignment would be just as infeasible.[10] Consider the idea of using source country control, precursor chemical regulations, and/or border interdiction to keep illegal drug prices high and thereby prevent drug use. Billions of dollars are invested in such strategies every year in the face of active, sometimes acrimonious debate about effectiveness, and there is no realistic possibility of resolving the debate via RCTs. Indeed, RCTs are often hard to achieve for *policies*, as opposed to *programs*, and even for programs when the unit of analysis is the city or larger, as opposed to the individual, classroom, or neighborhood.

Hence, analysts have developed a suite of quasi-experimental methods for assessing the extent to which a program or policy causes one or more outcomes when RCTs are not possible, including propensity score methods, instrumental variable techniques, difference-in-difference estimation, and regression discontinuity approaches. It is beyond the scope of this paper to describe them, but see Rosenbaum (2010) or Angrist and Pischke (2010) for a review of these methods, their strengths, and their limitations.

While quasi-experimental methods are generally viewed as second best, it is important to remind readers that even the gold standard RCT can also get tarnished in at least two ways. The first is external validity. Typical concerns pertain to selection effects associated with getting into the data set. A famous controversy pertains to the Life Skills Program. Effects in schools that implemented the program with high fidelity were impressive (Botvin, Baker, Dusenbury, Botvin, & Diaz, 1995), but some argue that an intent-to-treat rather than treatment-on-the-treated perspective would be more informative regarding economic returns that could be expected from future investments in the program (Gorman, 2002).

[10] Return on investment language, if not calculations, does appear in discussions of conflict prevention, as in this quote from *New York Times* columnist Nicholas Kristof (2011): "Sometimes foreign aid and diplomacy constitute a good investment, earning a strong return—and I would argue that that is the case with support for South Sudan. If South Sudan falls apart, we'll end up spending billions and billions of dollars in coming years and decades dealing with humanitarian and other catastrophes. Isn't it smarter—as well as simply the right thing to do—to help it stand on its feet in the next couple of years? Indeed, I would argue that the very modest sums we've invested over the last few years in Sudan diplomacy appear to have averted a new war, at least so far—and one estimate is that the average African war imposes economic costs of about $100 billion. Not a bad return."

We think it is also instructive to remember the case of the Cyclops light or, more formally, center high mounted stop lamps (CHMSLs), even though the original experiments were not RCTs (Kahane & Hertz, 1998). CHMSLs, or third brake lights, have been required on all passenger vehicles sold in the United States since 1986. Before the regulation was promulgated, three experiments were conducted with fleets of taxicabs and corporate cars; all yielded roughly 50 % reductions in rear-impact crashes in which the lead car was braking, which is roughly two thirds of rear-impact crashes. Yet the first study conducted after nationwide implementation found reductions of only 15 % and the news kept getting worse. A year later the estimated reduction was only 11.3 %. Within a decade, it had fallen to 5 %, a full order of magnitude smaller than initial studies found. There was nothing wrong with the early studies. However, before nationwide implementation, CHMSLs were novelties that attracted the attention of other drivers; 10 years later, people were so used to cars having third brake lights that the novelty effect had disappeared, leaving only the much smaller long-run effect. CHMSLs are so cheap that they remain a cost-effective way of preventing crashes, but the episode offers a poignant example of how even a perfect trial may vastly over- (or under-) estimate the true causal effects of a prevention intervention at scale.

A second limitation arises when effects on proximate outcomes are solidly grounded in RCTs, but economic analysis requires estimates of effects on other, longer run and/or more distal outcomes. As Caulkins (2002, p. 488) observes for school-based drug prevention:

> [T]here are clearly stated confidence intervals around empirical estimates of the programs' impact on participants' self-reported marijuana use through 12th grade. The real benefits, however, come from actual (not self-reported) changes in the use of all substances (not just marijuana) throughout the lifetime (not just 12th grade) of all people affected by the program, including spillover effects (not just reductions in use by people in the program). Uncertainty concerning how to extrapolate from the empirically measured evaluation outcomes to the outcomes of interest to a policy analyst dwarfs uncertainty concerning the magnitude of the measured effects.

That is, economic analysis sometimes requires projecting or modeling results for which suitable data have not been collected in the past. Modeling is common in engineering. Space scientists do not estimate how much fuel a rocket will need by collecting data on how much fuel rockets used in the past and running a regression; they build a model of the rocket's dynamics from first principles, such as Newton's laws of motion. Individual people are rarely as predictable as rockets, but sometimes they behave predictably in aggregate. For example, models of infectious diseases may help predict how certain interventions will affect the spread of hepatitis or HIV (Kaplan, 1995; Pollack, 2001, 2002).

Stepping back from rarified debates about what are and are not proper analytic approaches to establishing causality, there are also more pedestrian challenges. For example, federal prevention grants to states may be disbursed to a large number of agencies running a heterogeneous range of interventions. Each program might be so small that adding outcome evaluation would cost more than the program itself, but aggregating across programs may be impossible if the programs are unlike. This

sometimes leads to a strategy of merely verifying adherence to a model program approach versus directly assessing effects and implicitly assuming that the economic performance achieved by the documented model program applies to the present population. Such extrapolation is dangerous even if the model program were evaluated by a rigorous RCT.

Recognition of the limitations of individual trials—even RCTs—has given rise to the ideal that policy should be grounded in systematic reviews of multiple studies; examples include Cochrane reviews such as Foxcroft, Ireland, Lowe, and Breen (2002) or Thomas and Perera (2006). But even systematic reviews have limitations (Mullen & Ramírez, 2006); as McCambridge (2007) and others argue, reporting biases can influence what results get published in peer-reviewed journals. Furthermore, policy encompasses a broad range of questions and interventions, but systematic reviews tend to focus on effect size, not optimal design considerations, and the bulk of scientific evidence that meets strict review criteria may pertain to one of a few modalities. For drug prevention, evaluations of school-based programs predominate. Faggiano et al.'s (2008) Cochrane review of school-based programs included 32 studies versus only 17 for all nonschool modalities combined in Gates, McCambridge, Smith, and Foxcroft's (2006) review. Likewise, for harm prevention interventions, the literature on syringe exchange and safe injection facilities exceeds that of all other interventions combined (Ritter & Cameron, 2005).

All Consequences

Naturally economic analysis seeks to associate with a program all of its consequences. This effort entails an understanding of what and who is affected. Sometimes that is easy. When treating an individual's cancer, it is reasonable to presume that all of the resulting health benefits will accrue to the patient and, perhaps, to the patient's family. That circumscribes the range over which measurements need to be taken.

With prevention interventions, it is the norm not the exception for indirect or spillover effects to be too large to ignore. That is obvious with contagious disease. Preventing one person from becoming infected with HIV might avert more than one HIV infection if the individual receiving the prevention intervention would otherwise have infected others who in turn could have infected still others. In short, there is a "social multiplier" effect that a comprehensive analysis should recognize.

Including social multiplier and other spillover effects is not always easy. An epidemic's reproductive number varies not only by infectious agent but also by stage of the epidemic and the larger social context. And the challenges multiply when the physics of diffusion are less mechanistic. Not only literal pathogens but also drug use could be "contagious" in the sense of spreading from person to person (Brill & Hirose, 1969; Caulkins, 2005; Ferrence, 2001). Adoption of drug use can be seen as a form of new product adoption, something markets have studied with models very similar to susceptible-infected-recovered (SIR) epidemiological

models at least since the work of Bass (1969), and similar thinking applies to other behaviors, even urban legends (Noymer, 2001).

That social interactions can create indirect effects does not mean that it is easy to estimate them from the data. Great excitement surrounded the idea that obesity might be contagious (Christakis & Fowler, 2007), but great humility followed the discovery that similar analytic methods found acne, headaches, and height to be similarly contagious (Cohen-Cole & Fletcher, 2008).

Spillover effects can be even less direct. Blumstein and Cork (1996) make a convincing case that crack use spawned street markets that created "jobs" for youthful dealers, giving those adolescents the means and incentive to arm themselves. That arms race amplified the lethality of everyday disputes over slights and reputations and girlfriends that might previously have been settled with fists or knives. So it is entirely plausible that some demand-side intervention, whether prevention- or treatment-oriented, that successfully reduced use and consequently, drug-selling and thus arming and thus shooting, could have prevented some homicides. The precepts of economic analysis demand that all consequences be considered, but it is often not plausible to capture such indirect effects at all and certainly not with the same rigor as the direct effects experienced by program participants.

The technical jargon for this is *ambiguity regarding system boundaries*. Conceptually one wants to draw a dotted line around the system an intervention is targeting and measure all relevant outcomes for everything and everyone within those dotted lines. But prevention interventions, particularly drug and crime prevention interventions, are intervening in lives that are intertwined with others in the complex fabric of social interactions. The judgment inherent in deciding what effects can reasonably be estimated and which are so speculative and indirect as to require omission creates abundant opportunity for bias in evaluations, conscious or otherwise. Particularly given how often prevention interventions are evaluated by the entrepreneurial researchers who invented them, there is a temptation to include indirect benefits and exclude indirect costs.

Coda: The Goldilocks Principle for Economic Analysis in Prevention Sciences

We began with what appears to be an uncontroversial assertion. Allocation of taxpayers' dollars across programs should be informed by economic analysis that considers *all consequences* for society that are *caused by* a program, with options ranked based on some *common metric* that reflects the *present value* of cumulated benefit net of *all costs including opportunity costs*. We then marched through those five emphasized phrases pointing out how muddy this can become in practice, particularly in the context of prevention. Hence, when seeking the ideal role of cost analysis in prevention policy, we suggest turning to the wisdom of Aristotle and

Goldilocks. In the latter's spirit, we might think of three positions, albeit in stylized caricature.

In the bad old days prevention was not held accountable by cost analysis. We funded and researched and implemented on faith because "kids are worth it" and "an ounce of prevention is worth a pound of cure." Those were—and to the extent that the era hasn't fully passed still are—bad days for taxpayers. Lots of money and time have been squandered on programs whose performance just doesn't withstand scrutiny. But an even bigger misallocation may stem from funding programs that are effective, but relatively inefficient, because even more lives could have been saved by shifting those funds to programs that were still more cost-effective.

Times have changed. Policy domains that were once seen as beyond quantification are now routinely subjected to economic analysis, and those who still object have been relegated to a shunned fringe. The triumphant or barbaric sweep—depending on one's perspective—of economic analysis has moved from such traditional domains as defense, transportation, and flood control to tackle less quantifiable issues in environmental, health, and safety regulation and beyond, so by now any implicitly acknowledged exception for prevention is disappearing. Prevention investments are being subjected to the same brutal calculus as are other programs.

Is the pendulum swinging too far? Economic analysis is always harder in practice than it is on the chalkboard, but to the extent that the issues outlined above make its abstractions particularly problematic for prevention programs, one might say we have moved from the bad old days of no evaluation to the bad current days of rigidly formulaic evaluation.

We look forward with guarded optimism to a golden future of Aristotelian moderation. One that requires economic analysis of all prevention efforts, but which recognizes as even Arrow et al. (1996) do, that although benefit-cost analysis is useful, it is neither a necessary nor sufficient basis for public decision making. So, in their words (1996, pp. 221–222), "*Although agencies should be required to conduct benefit-cost analyses for major decisions and to explain why they have selected actions for which reliable evidence indicates that expected benefits are significantly less than expected costs, those agencies should not be bound by strict benefit-cost tests*. Factors other than aggregate economic benefits and costs, such as equity within and across generations, may be important in some decisions" [emphasis in original]. We would add that some license be given for adapting the methods to the peculiar exigencies of prevention. For example, varying the social discount rate within the range generally viewed as reasonable can decisively affect the outcome of analysis for prevention programs whose effects occur decades in the future.

Admittedly, this is a caricature. The idea that prevention should be subject to economic analysis is hardly new (cf. Plotnick, 1994). And even today the most widely implemented school-based drug prevention curriculum, D.A.R.E., is one for which evidence of effectiveness let alone cost-effectiveness is weak at best (Kanof, 2003). But we have a sense of pendulum swings and at this point worry as much about dogmatic faith in mechanical application of economic analysis as we do

about dogmatic faith in funding prevention without asking whether such spending is justified. And we look forward to a time when all prevention investments are accompanied by quantitative projection of costs and benefits produced and interpreted by people who are rooted in the prevention field and its special circumstances.

What might that involve? In part, just waiting. Scientists from a number of disciplines are already working hard to tackle many of the issues we identified above, in part because they are not unique to the prevention sciences. Methodological progress is particularly rapid for identifying causal relationships between programs and their many consequences, especially those intangible consequences such as preventing lost quality of life among dependents and their families. We can also expect progress on understanding the full range of costs, including the opportunity costs of volunteers and participants, and progress on how to value outcomes in comparable terms regardless of their form or when they occur.

It would be a mistake, though, to imagine the challenges are primarily technical, in the sense of just needing a new and fancier statistical method. The challenges are at least as much contextual and cultural; the field needs more people who are fully bilingual in economic analysis and prevention science.

We would also suggest that the greatest contributions will come when economic analysis is harnessed to produce insights, not just winners in a cost-effectiveness horse race. Classically economic analysis sells itself as promoting efficient use of scarce resources (e.g., taxpayer money) by determining which programs ought to be funded at all and which ought to be funded first. And in theory economic analysis can compare the benefits of prevention programs across target behaviors (e.g., comparing a drug prevention program to one that prevents obesity or promotes vaccination). Yet even that is still just picking winners from an existing stable. Economic analysis has the potential not only to identify which programs are winners but also *why they are winners*—in the sense of showing which benefits and costs are the most important, and which are secondary—and, hence, to provide insights that can inform the design of future generations of programs.

Acknowledgments John Carnavale, Rick Harwood, and Ted Miller made many useful suggestions that have improved this analysis.

References

Aldy, J. E., & Viscusi, W. K. (2008). Adjusting the value of a statistical life for age and cohort effects. *The Review of Economics and Statistics, 90*, 573–581.

Angrist, J. D., & Pischke, J.-S. (2010). *Mostly harmless econometrics*. Princeton, NJ: Princeton University Press.

Aos, S. (2010). *Return on taxpayer investment: Evidence-based options to improve statewide outcomes—update for the legislature*. Olympia, WA: Washington State Institute for Public Policy.

Aos, S., Leib, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*, Report #04-07-3901. Olympia, WA: Washington State Institute for Public Policy. Retrieved May 3, 2011, from http://www.wsipp.wa.gov/rptfiles/04-07-3901.pdf.

Arrow, K. J., Cropper, M. L., Eads, G. C., Hahn, R. W., Lave, L. B., Noll, R. G., et al. (1996). Is there a role for benefit-cost analysis in environmental, health, and safety regulation? *Science, 272*, 221–222.

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science, 15*, 215–227.

Blumstein, A., & Cork, D. (1996). Linking gun availability to youth gun violence. *Law and Contemporary Problems, 59*(1), 5–24.

Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2006). *Cost-benefit analysis: Concepts and practice* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

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.

Brill, H., & Hirose, T. (1969). The rise and fall of a methamphetamine epidemic: Japan 1945–55. *Seminars in Psychiatry, 1*, 179–194.

Caulkins, J. P. (2002). Using models that incorporate uncertainty. *Journal of Policy Analysis and Management, 21*, 486–491.

Caulkins, J. P. (2005). Models pertaining to how drug policy should vary over the course of an epidemic cycle. In B. Lindgren & M. Grossman (Eds.), *Substance use: Individual behavior, social interactions, markets, and politics*. Elsevier, Amsterdam: Advances in Health Economics and Health Services Research.

Caulkins, J. P., Everingham, S. S., Rydell, C. P., Chiesa, J., & Bushway, S. (1999). *An ounce of prevention, a pound of uncertainty, RAND MR-923-RWJ*. Santa Monica, CA: RAND Corporation.

Centers for Disease Control and Prevention. (2000). *Measuring healthy days*. Atlanta, GA: Centers for Disease Control and Prevention.

Christakis, N., & Fowler, J. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine, 357*, 370–379.

Cohen, M. A. (2005). *The costs of crime and justice*. New York: Routledge.

Cohen-Cole, E., & Fletcher, J. M. (2008). Detecting implausible social network effects in acne, height, and headaches: Longitudinal analysis. *British Medical Journal, 337*, a2533.

Cook, P. J. (2007). *Paying the tab: The economics of alcohol policy*. Princeton, NJ: Princeton University Press.

Daley, M., Shepard, D. S., & Bury-Maynard, D. (2005). Changes in quality of life for pregnant women in substance user treatment: Developing a quality of life index for the addictions. *Substance Use & Misuse, 40*, 375–394.

Drug Strategies. (1996). *Making the grade: A guide to school drug prevention programs*. Washington, DC: Drug Strategies.

Drummond, M. F., O'Brien, B., Stoddart, G. L., & Torrance, G. W. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). New York: Oxford University Press.

Ellickson, P. L., Collins, R. L., Bogart, L. M., Klein, D. J., & Taylor, S. L. (2005). Scope of HIV risk and co-occurring psychosocial health problems among young adults: Violence, victimization, and substance use. *Journal of Adolescent Health, 36*, 401–409.

Elvik, R. (2010). Strengthening incentives for efficient road safety policy priorities: The roles of cost-benefit analysis and road pricing. *Safety Science, 48*, 1189–1196.

Faggiano, F., Vigna-Taglianti, F. D., Versino, E., Zambon, A., Borraccino, A., & Lemma, P. (2008). School-based prevention for illicit drugs' use: A systematic review. *Preventive Medicine, 46*(5), 385–396.

Farrington, D. P., & Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology, 1*, 9–38.

Ferrence, R. (2001). Diffusion theory and drug use. *Addiction, 96*, 165–173.
Foxcroft, D. R., Ireland, D., Lowe, G., & Breen, R. (2002). Primary prevention for alcohol misuse in young people. *Cochrane Database System Review*, *3*.
French, M. T., Salome, H. J., Sindelar, J. L., & McClellan, A. T. (2002). Benefit-cost analysis of addiction treatment: Methodological guidelines and empirical application using the DATCAP and ASI. *Health Services Research, 37*, 433–455.
Gates, S., McCambridge, J., Smith, L. A., & Foxcroft, D. (2006). Interventions for prevention of drug use by young people delivered in non-school settings. *Cochrane Database of Systematic Reviews*, *1*. Art. No.: CD005030. doi: 10.1002/14651858.CD005030.pub2.
Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine*. Oxford, UK: Oxford University Press.
Gorman, D. M. (2002). "Science" of drug and alcohol prevention: The case of the randomized trial of the Life Skills Training program. *International Journal of Drug Policy, 13*, 21–26.
Greenwood, P. W., Model, K. E., Rydell, C. P., & Chiesa, J. (1998). *Diverting children from a life of crime: Measuring costs and benefits* (Revisedth ed.). Santa Monica, CA: RAND Corporation.
Hartmann, D., & Depro, D. (2006). Rethinking sports-based community crime prevention: A preliminary analysis of the relationship between midnight basketball and urban crime rates. *Journal of Sport and Social Issues, 30*, 180–196.
Harwood, H. (2000). *Updating estimates of the economic costs of alcohol abuse in the United States*: *Estimates*, *update methods*, *and data*. Report prepared by The Lewin Group for the National Institute on Alcohol Abuse and Alcoholism. Rockville, MD: National Institutes of Health.
Harwood, H., Fountain, D., & Livermore, G. (1998). *The economic costs of alcohol and drug abuse in the United States 1992*. Rockville, MD: National Institute on Drug Abuse.
Hawkins, D., Kosterman, R., Catalano, R., Hill, K., & Abbott, R. (2005). Promoting positive adult functioning through social development intervention in childhood: Long-term effects from the Seattle Social Development Project. *Archives of Pediatrics & Adolescent Medicine, 159*, 25–31.
Henderson, N., & Langford, I. (1998). Cross-disciplinary evidence for hyperbolic social discount rates. *Management Science, 44*, 1493–1500.
Humphreys, K., Wagner, T. H., & Gage, M. (2008, April 4). *Is it in medical centers' self-interest to provide substance use disorder treatment? A cost-consequence analysis in a national health care system*. Presented at the ISSDP Conference in Lisbon.
Johansson, P.-O. (2002). On the definition and age-dependency of the value of statistical life. *Journal of Risk and Uncertainty, 25*, 251–263.
Kahane, C. J., & Hertz, E. (1998). *The long-term effectiveness of center high mounted stop lamps in passenger cars and light trucks*, NHTSA Technical Report Number DOT HS 808 696. Washington, DC: National Highway Traffic Safety Administration. Retrieved December 29, 2010, from http://www.nhtsa.gov/cars/rules/regrev/evaluate/808696.html.
Kanof, M. E. (2003). *Youth illicit drug use prevention: DARE long-term evaluations and federal efforts to identify effective programs, GAO-03-172R*. Washington, DC: General Accounting Office.
Kaplan, E. H. (1995). A circulation theory of needle exchange. *Operations Research, 43*, 558–569.
Karp, L. (2005). Global warming and hyperbolic discounting. *Journal of Public Economics, 89*, 261–282.
Keeler, E., & Cretin, S. (1983). Discounting of life-saving and other nonmonetary effects. *Management Science, 29*, 300–306.
Kilmer, B., Caulkins, J. P., Pacula, R. L., MacCoun, R. J., & Reuter, P. H. (2010). *Altered state? Assessing how marijuana legalization in California could influence marijuana consumption and public budgets*. Santa Monica, CA: RAND Corporation. RAND Occasional Paper OP315.

Kind, P., Dolan, P., Gudex, C., & Williams, A. (1998). Variations in population health status: Results from a United Kingdom national questionnaire survey. *British Medical Journal, 316*, 736–741.

Kristof, N. (2011, January 11). Answering your Sudan questions, take 2. Retrieved January 13, 2011, from http://kristof.blogs.nytimes.com/2011/01/11/answering-your-sudan-questions-take-2/?scp=1&sq=sudan%20carter%20#2&st=cse.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*, 443–477.

Lee, S., Aos, S., & Miller, M. (2008). *Evidence-based programs to prevent children from entering and remaining in the child welfare system: Benefits and costs for Washington*, Document No. 08-07-3901. Olympia, WA: Washington State Institute for Public Policy. Retrieved January 13, 2011, from http://www.wsipp.wa.gov/rptfiles/08-07-3901.pdf.

McCambridge, J. (2007). A case study of publication bias in an influential series of reviews of drug education. *Drug and Alcohol Review, 26*, 463–468.

Mensch, B. S., & Kandel, D. B. (1988). Dropping out of high school and drug involvement. *Sociology of Education, 61*, 95–113.

Miller, T. R., Cohen, M. A., & Wiersema, B. (1996). *Victim costs and consequences: A new look, NCJ 155282*. Washington, DC: National Institute of Justice.

Miller, T. R., Romano, E. O., & Spicer, R. S. (2000). The cost of childhood unintentional injuries and the value of prevention. *Future of Children, 10*, 137–163.

Moore, M., Boardman, A. E., Vining, A. R., Weiner, K. L., & Greenburg, D. H. (2004). Just give me a number! Practical values for the social discount rate. *Journal of Policy Analysis and Management, 23*, 789–812.

Mullen, P. D., & Ramírez, G. (2006). The promise and pitfalls of systematic reviews. *Annual Reviews of Public Health, 27*, 81–102.

Newhouse, J. (1993). *Free for all? Lessons from the RAND Health Insurance Experiment*. Cambridge, MA: Harvard University Press.

Nocetti, D., Jouini, E., & Napp, C. (2008). Properties of the social discount rate in a benthamite framework with heterogeneous degrees of impatience. *Management Science, 54*, 1822–1826.

Noymer, A. (2001). The transmission and persistence of "urban legends": Sociological application of age-structured epidemic models. *Journal of Mathematical Sociology, 25*, 299–323.

Office of National Drug Control Policy. (2004). *The economic costs of drug abuse in the United States*. Washington, DC: Executive Office of the President. Publication no. 207303.

Plotnick, R. D. (1994). Applying benefit-cost analysis to substance use prevention programs. *The International Journal of the Addictions, 29*, 339–359.

Pollack, H. (2002). Methadone maintenance as HIV prevention: Cost-effectiveness analysis. In E. H. Kaplan & R. Brookmeyer (Eds.), *Quantitative evaluation of HIV prevention programs*. New Haven, CT: Yale University Press.

Pollack, H. A. (2001). Can we protect drug users from hepatitis C? *Journal of Policy Analysis and Management, 20*, 358–364.

Pyne, J., French, M., McCollister, M., Tripathi, K., Rapp, S., Booth, R., et al. (2008). Preference-weighted health-related quality of life measures and substance use disorder severity. *Addiction, 103*, 1320–1329.

Rabin, R., & de Charro, F. (2001). EQ-5D: A measure of health status from the EuroQol Group. *Annuals of Medicine, 33*, 337–343.

Rice, D. P., & Cooper, B. S. (1967). The economic value of human life. *American Journal of Public Health, 57*(11), 1954–1966.

Ritter, A., & Cameron, J. (2005). *A systematic review of harm reduction, DPMP Monograph Series No 06*. Fitzroy, Australia: Turning Point Alcohol and Drug Centre.

Rosenbaum, P. R. (2010). *Design of observational studies*. New York: Springer.

Schelling, T. (1968). The life you save may be your own. In S. Chase (Ed.), *Problems in public expenditure analysis* (pp. 127–162). Washington, DC: Brookings Institution.

Shepard, D. S., & Zeckhauser, R. J. (1984). Survival versus consumption. *Management Science, 30*, 423–439.

Stouthard, M. E. A., Essink-Bot, M. L., Bonsel, G. J., et al. (1997). *Disability weights for diseases in the Netherlands*. Rotterdam, Netherlands: Erasmus University.

Thomas, R., & Perera, R. (2006). School-based programmes for preventing smoking. *Cochrane Database System Review*, *3*.

U.S. Department of Health and Human Services, Administration for Children and Families. (2010). *Head Start impact study: Final report*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

U.S. Department of Housing and Urban Development, Office of Policy Development and Research. (2003). *Moving to opportunity for fair housing demonstration program: Interim report*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Viscusi, K. W., & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty, 27*, 5–76.

Weissberg, R. P., Kumpfer, K. L., & Seligman, M. E. P. (2003). Prevention that works for children and youth. *American Psychologist, 58*, 425–432.

Chapter 26
Strengthening Prevention Science to Ensure Effectiveness of Intervention in Practice: Setting up an International Agenda

Fabrizio Faggiano, Fabrizia Giannotta, and Elias Allara

Introduction

Primary prevention is the public health activity dedicated to prevent the occurrence of disease and promote health (Last, 2006). It can act by reducing exposure to well-known risk factors at the environmental level (e.g., air pollution, occupational exposure to chemicals, asbestos ban) by increasing individual biological defences (such as immunization) or act at the individual level to change risky behaviors.

There are deep differences in these three approaches to primary prevention: the first mainly addresses eliminating risk factors from the environment through regulatory interventions at a national or international level. The second and the third work on different and complex individual mechanisms. This contribution focuses on the third approach: the prevention and change of risky behaviors.

Psychological and Social Mechanisms of Prevention of Risky Behaviors

Interventions focused on the third approach usually target psychological and social factors that are recognized to be related to risk behaviors; such factors are usually the mechanisms through which interventions are designed to modify problematic behaviors. For example, parenting behaviors, such as communication, setting rules, and providing warmth and emotional support, have an inverse association to the development of high-risk behaviors of adolescents, such as alcohol and substance

F. Faggiano (✉)
Department of Translational Medicine, Avogadro University of Eastern Piedmont, Novara 28100, Italy
e-mail: fabbrizio.faggiano@med.unipmn.it

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2_26,

use, delinquency, and school dropout (Velleman, Templeton, & Copello, 2005). On the basis of these associations, interventions that intend to reduce substance use may want to focus on changing the type of communication and relationship between parents and children, and this in turn would reduce the risk of problematic outcomes during adolescence. Personal competencies, such as decision making, self-control, self-reinforcement, substance-specific assertiveness, attitude against the use of alcohol and drugs, and normative expectations, also are negatively associated with drug use during adolescence (Botvin, Malgady, Griffin, Scheier, & Epstein, 1998; Griffin, Botvin, Scheier, Epstein, & Doyle, 2002; Willis, Baker, & Botvin, 1989). Thus, many interventions have focused on the enhancement of such competences to diminish the occurrence of engagement in high-risk behaviors.

Unexpected Effects of Prevention Programs and Possible Explanations

In spite of the existence of clear theories about what are the psychological processes that should be targeted to make interventions work, an appropriate translation of this knowledge into practice cannot always ensure the success of the program. There are several examples of theory-based intervention programs that produced counterintuitive effects. Among those that adopted rigorous methods, the US National Youth Anti-Drug Media Campaign during the years 1998–2004 and the "Take Charge of Your Life" program for drug, tobacco, and alcohol prevention both were based on social influence theory, which is shown to be fundamental to most effective interventions (Tobler, 2000). Nevertheless, they produced, respectively, a substantial increase in the use of cannabis among adolescents exposed to television advertising (Hornik, Jacobsohn, Orwin, Piesse, & Kalton, 2008) and an increase in the risk of tobacco and alcohol use and binge drinking (Sloboda et al., 2009) among those who had not used tobacco or alcohol at baseline, when they were 12–13 years old.

Outside of the field of substance prevention, the examples are similar. The Spark study suggested that an intervention for the promotion of physical activity in elementary schools, aimed at reducing the risk of obesity, significantly increased body mass index among students of both intervention groups compared to controls 18 months after the intervention (Sallis et al., 1993). The Postponing Sexual Involvement curriculum, widely implemented in the USA and evaluated in California by a randomized study, determined the increase of intercourse, the number of sexual partners, and the number of sexually transmitted diseases (nonsignificant) as well as the number of of pregnancies ($P < 0.05$) among intervention groups, both of which were led by peers and adults (Kirby, Korpi, Barth, & Cagampang, 1997).

There are several possible reasons for these failures, ranging from theory to implementation. In the following, we attempt to group these reasons into five broad categories.

Bad Theory

Theories in psychological and sociological fields usually are based on longitudinal studies that attempt to establish temporal associations between factors even though they cannot claim causality. Sometimes interventions are grounded on theories that do not have another empirical base. For instance, one of the first interventions to reduce substance use among youth was aimed at increasing knowledge about substance use. The underlying theoretical framework was the information-dissemination theory that postulates a decrease of substance use as a consequence of the acquisition of knowledge about negative effects of drug use. However, it was found to be ineffective in experimental studies even though it was, in fact, able to increase knowledge of the consequences of substance use (Botvin, 1990).

Experimental evaluation studies can be considered as experiments that can test the validity of a theory. As claimed by Claude Bernard (1865) 150 years ago in his *Introduction to Experimental Medicine*, "when we meet a fact which contradicts a prevailing theory, we must accept the fact and abandon the theory, even when the theory is supported by great names and generally accepted."

This case of program failure is quite common in different contexts; for example, programs that involved organized visits to prison facilities by juvenile delinquents or children at risk of becoming delinquent and were designed to deter participants from committing future offenses by providing first-hand observations of prison life failed to prove their effectiveness. A systematic review of seven randomized trials found that this approach not only fails to deter crime but also actually leads to more offending behavior (Petrosino, Turpin-Petrosino, & Buehler, 2004), with an odds ratio of 1.68 (95 % confidence interval, 1.20–2.36).

Intervention Is Unable to Affect Target-Mediating Factors

Sometimes the proposed activities might be inappropriate to target and modify the factors associated with the risk behaviors, namely the mediating factors of the intervention (e.g., psychological factors that explain the risk behavior). When a lack of effect on mediating factors occurs together with a lack of effect on the targeted behaviors, one can argue that the intervention does not include the appropriate activities necessary to target the appropriate factors.

Intervention Targets the Wrong Mediating Factors

Sometimes the intervention can indeed modify the targeted mediating factors, but mediation analysis can reveal that those factors are not, in fact, mediators of the final outcomes. For example, Botvin, Griffin, Diaz, and Ifill-Williams (2001), in

their school-based intervention to prevent drug use, found effects on many expected mediators. However, only a few of them were found to be related to the change of the outcome (e.g., substance use).

Implementation Failure

A program also may fail because of poor program implementation. Interventions are usually constituted of components (e.g., a series of role-plays that simulate real-life situations) that are aimed at modifying specific factors (e.g., refusal skills), which in turn are expected to change the final outcome. When implementation failure occurs, it is likely that the program components were not correctly implemented. For example, it might have not been possible for some teachers to deliver all of the program contents. If a program is not implemented as intended, no conclusion can be drawn about its theoretical validity and efficacy.

Good Theory but Sensitive to the Context

Finally, sometimes moderator mechanisms may influence the relationship between mediating factors and the target behavior, resulting in an intervention's lack of effect. In other words, psychological factors may interact with many other social or contextual factors and may result in negative outcomes. For instance, peer-training programs, even when they target well-established mediators, obtained iatrogenic effects if implemented among a group of high-risk youth (Dishion, McCord, & Poulin, 1999). As Dishon and colleagues (1999) argued, high-risk youth in the group influenced each other into a "deviancy training" that resulted in an increase in deviant behaviors after the intervention. This suggests that even when an intervention targets the right mediators, moderator factors might intervene and cause null or iatrogenic effects.

In sum, the complexity of psychological pathways and their interaction with context may complicate the identification of the right mediators to target and the activities that are aimed at modifying those mediators. Rigorous scientific standards for the evaluation of prevention interventions should include a continued process of assessment of the robustness of the theory and of causal relationships, taking into account psychological and social processes that may act as moderators.

Need for Evaluation of Prevention Interventions

In the previous sections we discussed the fact that prevention interventions target complex and delicate psychosocial mechanisms that are expected to be determinants of problematic behaviors. However, we also showed that preventive

interventions provide unique opportunities to test empirically theories of human behavior through experimental designs. This calls for a rigorous evaluation of both the intervention effects and the mediator and moderator mechanisms that are responsible for those effects. The need for a rigorous evaluation process is shared with other fields that deal with individual well-being and health as well, for example, pharmacology, but with deep differences in implementation. As David Sackett (2002) wrote a few years ago in his polemic article "... surely the fundamental promise we make when we actively solicit individuals and exhort them to accept preventive interventions must be that, on average, they will be the better for it. Accordingly, the presumption that justifies the aggressive assertiveness with which we go after the unsuspecting healthy must be based on the highest level of randomized evidence that our preventive manoeuvre will, in fact, do more good than harm. Without evidence from positive randomized trials (and, better still, systematic reviews of randomized trials) we cannot justify soliciting the well to accept any personal health intervention."

In spite of these premises, it is highly likely that the great majority of prevention activities provided in practice have never been evaluated. For example, according to the 2013 report of the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), the effectiveness of prevention programmes is rarely evaluated in Europe, although "A small shift has [..] been noted towards the use of positively evaluated universal prevention approaches in schools" (EMCDDA, 2013). Most European prevention interventions are built upon untested theoretical assumptions and individual experience/perception, and all European countries have experienced a low level of evaluation activity, with few exceptions, according to the opinion of an expert panel (EMCDDA, 2014a). In Italy, for example, the proportion of evidence-based prevention interventions targeting youth for the prevention of tobacco smoking, alcohol abuse, and the promotion of a healthy diet and safe sexual activity is probably less than 1 % (Coffano, 2009). Even if in the USA evaluation of effectiveness of prevention interventions seem to be more common (Gandhi, Murphy-Graham, Petrosino, Chrismer, & Weiss, 2007), prevention is far from the standards of many others field of medicine.

The situation is even worse when it comes to understanding why those interventions are effective. Indeed, there are a very limited number of studies that show mediating mechanisms, as Sandler, Schoenfelder, Schoenfelder, and MacKinnon (2011) observed in a review of mediators of family-based interventions. This gap in our knowledge translates into a failure in our understanding of why and how interventions work and thus difficulty in transferring the effective intervention to other contexts.

International Procedures to Disseminate Evidence-Based Prevention Programs

The diffusion of effective, evidence-based practices in the field of prevention is still limited. However, a formal process aimed at restricting the use of interventions to those with evidence of effectiveness is currently lacking. No procedures are currently in place for prevention interventions that are similar to those used in other health areas, such as the approval of medications used by the Food and Drug Administration (FDA; http://www.fda.gov) and the European Medicines Agency (http://www.ema.europa.eu).

Nevertheless, some interesting experiences are worth mentioning. In the field of mental health and substance abuse, for example, the EMCDDA established the Best Practice Portal (BPP) to respond to the 2009–2012 EU drugs action plan (EMCDDA, 2014b). The BPP summarizes evidence gathered from systematic reviews (mainly the Cochrane Library) assessing the effectiveness of drug-related prevention interventions. The BPP contains available evidence regarding specific activities or interventions aiming to prevent and treat drug abuse and reduce harm caused by drug abuse (EMCDDA, 2014b). The BPP includes only general programs categories instead of specific program names (e.g., "comprehensive family-oriented prevention for cannabis use"), scored with the GRADE evidence assessment system (Guyatt et al., 2008) to determine the level of impact on patients, from "beneficial" to "evidence of ineffectiveness" (Bo, Allara, & Ferri, 2011).

In the USA, comprehensive best-practice lists of evidence-based prevention programs are regularly updated by governmental or academic agencies, particularly in the fields of mental health, drug abuse, and violence (Gandhi et al., 2007). Prominent lists are the List of Exemplary and Promising Prevention Programs, run by the US Department of Education, in which interventions are selected using criteria such as evidence of efficacy, quality of the program, educational significance, and usefulness to others (http://www2.ed.gov/admins/lead/safety/exemplary01/exemplary01.pdf); the National Registry of Evidence-based Programs and Practices, run by the Substance Abuse & Mental Health Services Administration, lists programs that are reviewed and rated by independent reviewers, describing them in light of the quality of the evaluation research and readiness for dissemination (http://nrepp.samhsa.gov/).

We were able to find only a small number of similarly advanced experiences in other fields of prevention. The US Centers for Disease Control and Prevention, for example, runs the Compendium of Evidence-based HIV Prevention Interventions (http://www.cdc.gov/hiv/topics/research/prs/evidence-based-interventions.htm) that brings "science-based, community, group, and individual-level HIV prevention interventions to community-based service providers and state and local health departments." It presents best evidence interventions and promising-evidence interventions, classified according to the evaluation quality and effect size of interventions. The California Evidence-Based Clearinghouse for Child Welfare (http://www.cebc4cw.org/) provides interventions to prevent mental health disorders such as depression, anxiety, and neglect for adults, children, and families.

The classification of interventions is based on the level of research support available. The Office of Juvenile Justice and Delinquency Prevention (Model Program Guide contains interventions designed to prevent at-risk behaviors spanning from "academic problems" to "aggression/violence" and "gang activity." A similar project is carried out by the University of Colorado's Blueprints for Violence Prevention (Center for the Study and Prevention of Violence, Institute of Behavioral Science; http://www.colorado.edu/cspv/blueprints/).

We also found one lists of studies: the Washington State Planning Group's Effective Intervention and Strategies Document (http://www.doh.wa.gov/YouandYourFamily/IllnessandDisease/HIVAIDS/Prevention/Interventions.aspx), that annually review and update a list of effective HIV prevention programs.

Unfortunately the lists vary as to which programs are classified as effective and which are not. And, at least in the field of drug abuse, the lists vary considerably as to the evidence needed for inclusion, the use of independent assessors, the required length of follow-up duration, as well as issues such as outcome reporting bias and generalizability of programs (Gandhi et al., 2007). Table 26.1 compares the criteria for program classification adopted by the seven lists reviewed by Gandhi (2007) together with those of the EMCDDA's BPP (2014b).

In this panorama, characterized by large variability in the definition of quality standards, the work of Flay et al. (2005) deserves special mention. On behalf of the Society for Prevention Research, standards of evidence for prevention interventions were developed that should be considered a reference for reviews of prevention programming. Three different sets of standards are described: (1) efficacy, the utility of the intervention under ideal conditions; (2) effectiveness, the efficacy under real-world conditions; and (3) readiness for dissemination. The article presents a detailed list of standards for all three situations regarding the quality of outcomes, measurements, study design, and statistical analysis; the establishment of minimum duration of effects; and the role of replication, generalizability requirements, and related issues.

Despite the effort of these agencies, especially from the USA, a unique, harmonized, mutually accepted international system to process and classify evidence on prevention and that is available to decision makers is still lacking. Such a system could (a) facilitate adoption of effective, evidence-based interventions; (b) reduce dissemination of untested and harmful interventions; and (c) assist decision makers in dispersing public and private money more efficiently.

Intrinsic Limits of Current Procedures for the Improvement of Evidence-Based Prevention: Ingredients Versus Programs

Although useful, the procedures and systems described earlier do not ensure an adequate level of dissemination of effective interventions nor a reduction in the delivery of ineffective ones. This is mainly because of the intrinsic characteristics of the procedures. First, they suggest that programs be adopted voluntarily, without

Table 26.1 Criteria used in evidence-based lists to classify prevention interventions. This table is partly based on work by Gandhi et al. (2007), with the addition of BPP EMCDDA (2014b)

List	Criteria							
	Evidence of efficacy	Quality of evaluation	Quality of program goals	Quality of program rationale	Quality of program content and appropriateness	Quality of program implementation methods	Educational significance[a]	Usefulness/ replicability
Blueprints	Yes	Yes	Yes	N/A	N/A	N/A	N/A	Yes
Drug strategies: making the grade	Yes	Yes	Yes	N/A	N/A	N/A	N/A	N/A
ED list	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Maryland report	Yes	Yes	N/A	N/A	N/A	N/A	N/A	N/A
NIDA guide	Yes	Yes	N/A	N/A	N/A	N/A	N/A	N/A
SAMHSA National Registry of Evidence-Based Programs and Practices	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Youth violence: a report of the surgeon general	Yes	Yes	Yes	N/A	N/A	N/A	N/A	Yes
EMCDDA best practice portal	Yes	Yes	Yes	No	Yes	Yes	No	No

EMCDDA European Monitoring Center for Drugs and Drug Abuse, *NIDA* National Institute on Drug Abuse, *SAMHSA* Substance Abuse and Mental Health Services Administration, *ED* US Department of Education, *N/A* not available

[a]The application describes how the program is integrated into schools' educational missions

any incentive or disincentive. Their use relies mainly on the authoritativeness of the agency that maintains the lists, and this quality varies widely across countries, regions, and probably social conditions.

A further limitation undermines the above-described procedures: the oversimplification of program evaluation. In a recent editorial, Heneghan noticed, disappointingly, that the current evidence of effectiveness of counseling and education to change behavior in order to prevent cardiovascular diseases does not show an effect in general populations. One of the reasons put forward to explain this observation is that considerable variation often occurs in the components of the interventions (Heneghan, 2011). This is a basic characteristic: interventions aimed at changing risky behaviors are usually complex and have multiple components. In a recent review of the content of effective substance abuse prevention programs, Hansen, Dusenbury, Bishop, and Derzon (2007) identified 23 distinct content areas in 48 programs. The average program addressed 8.5 content areas. Moreover, they noticed that most programs are an amalgam of approaches and content areas that are independent of formal theories. The review also showed that programs are not truly driven by theory and do not adhere consistently to a theory's tenets. On the other hand, Noar and Zimmerman (2005) demonstrated that there is a large overlap among theories commonly used as the basis for interventions. To our knowledge, similar analyses of prevention program content are not common in other fields of prevention, but, presumably, the above reasoning can be extended to those domains.

Thus, disappointingly, what we know about the effect of prevention programs is just the total effect, and it is impossible to disentangle the role of specific components or theories. This makes it impossible to isolate effective components from those programs that are ineffective or even iatrogenic, even if this could be highly important for the elaboration of new programs.

Examples of Frameworks for the Evaluation and Approval of Complex Interventions

A framework that is useful for explicitly addressing the evaluation of complex interventions (MRC, 2000; Campbell et al., 2000) outlines four phases that would help define the "active ingredients" of prevention interventions. In addition to the "preclinical" or theoretical phase, which must frame the conjectural basis of the effectiveness of the intervention, the four phases include the following: (a) *Phase I, or modeling*: This phase involves delineating an intervention's components and how they interrelate and how active components of a complex package may relate to either surrogate or final outcomes. It may also include qualitative testing through focus groups, preliminary surveys, case studies, or small observational studies. (b) *Phase II, or exploratory trial*: In this phase the evidence gathered thus far is put to the test, with the possible use of experimental designs, varying different components to see what effect each has on the intervention as a whole. This phase

could permit the testing of alternative forms ("doses") of an intervention. (c) *Phase III, or main trial*: This involves a randomized controlled trial to evaluate the main effect a complex intervention. (d) *Phase IV, or long term surveillance*: This phase consists of a separate study to establish the long-term and real-life effectiveness of the intervention. The broader applicability of an intervention outside of a research context may be tested, and rare or long-term adverse events are identified. This stage is likely to involve observational designs.

This framework addressed many issues that seem to be essential to prevention interventions. For example, the preclinical phase focuses on the identification of the "active ingredients" of the intervention; besides the component of the intervention, included in this concept are all the other factors that can affect its effectiveness, such as the characteristics of the patient, the skill level of providers, the method of delivery, frequency, duration, and timing. Although the exploratory phase is explicitly targeted to "consider variants of the intervention and their possible effects on outcomes," it fails to answer the main question: Which ingredients work and which do not work?

Several years after this framework was introduced, Collins, Chakraborty, Murphy, and Strecher (2009) suggested a way to tackle the challenge of assessing the effects of single program components and their combinations. They proposed adopting a phased experimental approach – essentially a randomized trial using a factorial design to test separately each active component identified in the previous phases of the process. Factorial designs can clearly help to evaluate separately several components and identify possible interactions. However, it cannot be the solution in the field of prevention because the number of active ingredients to be tested is overwhelming; following the review by Hansen et al. (2007), an average number of 8.5 components per intervention have to be tested across, for example, different dosages and methods of delivery.

The oldest and most comprehensive framework for developing and evaluating health technologies is one dedicated to the approval of medications. In both the USA and Europe there are official agencies responsible for the authorization and licensing of drugs: the US FDA and the European Medicine Agency (EMA), respectively (Fig. 26.1). Although formal approval pathways are slightly different for the two agencies, the processes are similar. Once biochemical research indentifies a new candidate molecular target and screens thousands of chemical compounds to generate promising molecules with desirable properties, four steps have to be accomplished to apply for approval of a new drug: preclinical testing and three clinical trials phases (Pharma, 2011). Preclinical testing consists of laboratory (in vitro) and animal (in vivo) testing performed to determine the drug's safety profile (Pharma, 2007). Clinical trials then determine whether a potential drug can safely and effectively treat a targeted disease in humans and consist of three phases. Phase 1 involves 20–100 healthy volunteers and aims to study the drug's pharmacokinetics and pharmacodynamics to determine the dosage and safety of the candidate new drug. This phase is similar to the preclinical testing but involves humans and not laboratory studies or animal testing. Phase 2 involves 100–300 patient volunteers to be studied for about 2 years to evaluate the drug's effectiveness and short-term side effects. Finally, during Phase 3, 1,000–3,000 patient

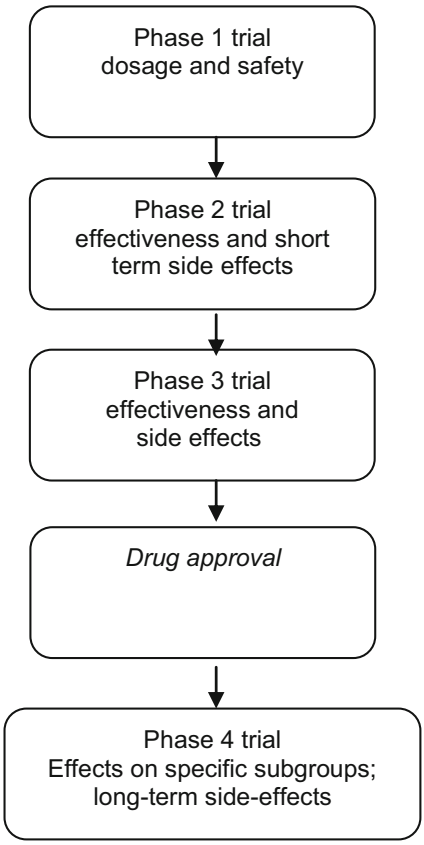

Fig. 26.1 The process of approval of drug medicines adopted by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA)

volunteers are involved to verify the drug's effectiveness and monitor adverse long-term effects. Phase 3 trials are both the costliest and longest trials, and they sometimes involve hundreds of sites in the around the world. After the third phase is completed, the drug manufacturer applies for EMA and/or FDA approval. In Europe drug licensing may happen through a centralized or a decentralized (or "mutual recognition") system (http://www.ema.europa.eu).

When a drug receives marketing authorization, the manufacturer may apply to market its drug in each individual country. After this last step, the product is available for doctors to prescribe.

In the USA, after completing clinical trials, the company sends the FDA a New Drug Application, which contains all the preclinical and clinical information obtained during the testing phase (http://www.fda.gov).

Once a drug is approved, it can be marketed (Lipsky & Sharp, 2011). Where the product is marketed and sold, and as a much larger number of patients begin to use the drug, drug manufacturers are required to monitor drug safety and submit periodic reports, including cases of adverse events, to the FDA and/or EMA. In addition, a company may be required to conduct additional studies on an approved

drug in "Phase 4" trials: these studies can be set up to evaluate long-term safety or how the new medication affects a specific subgroup of patients.

The Medical Research Council (2000) framework acknowledges parallels in the sequence of steps usually required in the evaluation of drugs from initial preclinical research through surveillance after marketing. However, it also stresses that such parallels and analogies in no way imply that the evaluation of prevention programs is like that of a new drug because of the numerous differences and levels of complexity.

International Perspective and the Role of International Societies

In the USA the establishment of the Society for Prevention Research (SPR) has had a remarkable impact on the development of the field of prevention science. Founded in 1991, the SPR's scope has broadened over the years, together with an increase in the number of its collaborations with varied federal US funding agencies such as the National Institute on Drug Abuse, the National Institute of Mental Health, and others (http://www.preventionresearch.org/about-spr). In 2010, the European Society for Prevention Research was founded. Among its purposes is the promotion of high-quality intervention assessment studies in Europe (http://www.euspr.org). A similar experience is being led in the Middle East and is forming the basis for the establishment of a new pan-Arab agency for drugs monitoring and prevention, the Arab Research, Resource and Information Center on Drug Prevention.

There are other international experiences that provide a glimpse into the potential of international collaboration in the field of prevention. The US-EU Cross-National Study on Variation of Receptivity to Substance Abuse Prevention Interventions is an opportunity to explore international differences and factors affecting program effectiveness focusing on the EU-Drug Abuse Prevention study (EU-DAP) and the Adolescent Substance Abuse Prevention Study (ASAPS). Both of these studies consist of large longitudinal, randomized, controlled trials conducted in Europe and in the USA, respectively, assessing comparable preventive interventions implemented among similarly aged populations. The current in-depth analysis being carried out by both study groups will contribute to an understanding of the role that program context plays in determining the effect of an intervention (Brown et al., 2013).

Another example is ALICE RAP (Addiction and Lifestyles in Contemporary Europe—Reframing Addictions Project): this new dynamic transdisciplinary EU project aims to help policy makers to "re-think and re-shape" current and future approaches to the huge human and economic costs of addictions and risky lifestyles in Europe. ALICE RAP aims to critically examine and analyze currently fragmented research and strengthen scientific evidence to inform a new dynamic

platform for public and political debate on current and alternative approaches to addictions (http://www.alicerap.eu).

The vitality of the research arena gives hope as to the role that research will play in strengthening the evidence base for prevention; in particular, the capacity of research networks to work and think together, share objectives, and study results, such as the ALICE RAP and SPR experiences, is a requirement for any new advancement in prevention science.

A Proposal for a Four-Pillar System of Approval of Intervention Programs

In the previous sections we argued that there are basic conditions to moving forward in the prevention field: fostering prevention science and raising the field's scientific standards. This should include, and be hastened by, setting up a formal process for the approval of prevention interventions analogous to that of medicines. This process, based on evidence of effectiveness and safety, should have some peculiarities and be based on the following pillars:

1. *A formal process of approval*. The process should be administered by agencies recognized at an international level. The whole process should lead to the release of a certification of effectiveness and safety. Single countries could give priority to certificated programs in health planning or could, once the number of certified programs is sufficient to cover all the needs of prevention, consider making certification compulsory to program delivery.
2. *A three step evaluation process*. (1) Measure the effect of any single intervention component on specified/related mediators of the target behavior through an experimental study. This would allow shorter follow-ups and smaller sample sizes. This step should also address the effective dose through testing, when appropriate, different frequencies or durations of interventions, as well as testing combinations of interventions. For single-component programs, the evaluation should be limited to effective dose evaluation. (2) Conduct a randomized controlled trial with a rigorous design and appropriate sample size to evaluate the full program in an appropriate and generalizable population. As an alternative, a meta-analysis of smaller trials could also be considered. After this step the certificate could be released. (3) Require replicability, in particular in different contexts, to assess the generalizability of results and the safety of the intervention. Replication of the original program or of modified versions is always recommended, especially if conducted by different groups of researchers (Valentine et al., 2011). Evaluating versions adapted for other languages or cultures also is recommended and should be submitted to the process to upgrade the certification (Fig. 26.2).
3. *Transparency and publication of all documents*. This aspect is of particular relevance for the documentation of the program itself and for its dissemination.

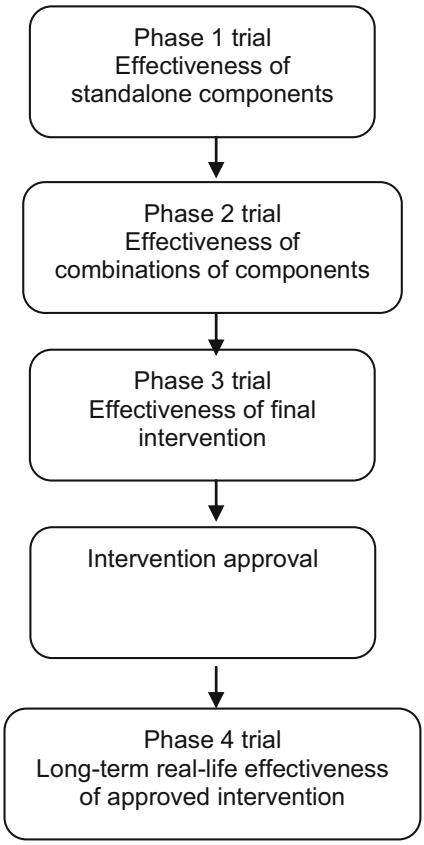

Fig. 26.2 Proposal for an approval process of prevention interventions

The level of documentation must be such that it enables replication with a high level of fidelity. The evaluation of single components is particularly important to allow other program developers to use effective components to build new programs from already validated components.

4. *International cooperation*. The establishment of standards of evidence for the approval of prevention programs requires international cooperation and consensus to develop a process accepted by researchers and practitioners. Large consensus and a high level of adherence, especially from developers and researchers, is needed since the impact of such a process on prevention practice can only occur once the number of programs and their ingredients can satisfy practitioners. This is one of the main challenges of this presented proposal.

Conclusions

The scientific content of prevention interventions and programs are still underconsidered and underused. Despite many commendable experiences, systems currently adopted by public agencies and scientific societies to ensure the effectiveness of prevention interventions adopted in clinical practice do not seem to have reached their principal aim, and a large part of prevention interventions found in practice, at least in Europe, do not have the minimal requirements to ensure that disseminated programs are effective, or at least safe.

Prevention scientists and professionals have to foster prevention science to promote the dissemination of effective intervention and to develop advanced methods for the evaluation of mechanisms and effects of prevention interventions. Among the possible items on the agenda of prevention science, we propose international cooperation to design and carry out a system of approval of prevention interventions that are able to target this objective. This chapter summarizes some proposals already presented by the scientific community and delineates the characteristics of a possible process used at an international level.

References

Bernard, C. (1865). *Introduction à l'étude de la medicine expérimentale*. J. B. Baillière et fils, Paris.

Bo, A., Allara, E., & Ferri M. (2011). Scientific evidence and practice: Bridging the gap. A European Monitoring Center for Drugs and Drug Addiction (EMCDDA) project to promote Best Practice in Drug Addiction field. In: *19th Cochrane Colloquium*. Madrid.

Botvin, G. J. (1990). Substance abuse prevention: Theory, practice, and effectiveness. *Crime and Justice, 13*, 461–519.

Botvin, G. J., Griffin, K. W., Diaz, T., & Ifill-Williams, M. (2001). Drug abuse prevention among minority adolescents: Posttest and one-year follow-up of a school-based preventive intervention. *Prevention Science, 2*(1), 1–13.

Botvin, G. J., Malgady, R. G., Griffin, K. W., Scheier, L. M., & Epstein, J. A. (1998). Alcohol and marijuana use among rural youth: Interaction of social and intrapersonal influences. *Addictive Behaviors, 23*(3), 379–387.

Brown, C. H., Sloboda, Z., Faggiano, F., Teasdale, B., Keller, F., Burkhart, G., et al. (2013). Methods for synthesizing findings on moderation effects across multiple randomized trials. *Prevention Science, 14*(2), 144–156.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., et al. (2000). Framework for the design and evaluation of complex interventions to improve health. *BMJ, 321*, 694–696.

Coffano, E. (2009). Guadagnare salute in adolescenza: Ricognizione delle esperienze di prevenzione e promozione della salute in Italia. Retrieved March 8, 2013, from http://www.dors.it/public/ar3601/REPORT_progettoAdolescenti.pdf.

Collins, L. M., Chakraborty, B., Murphy, S. A., & Strecher, V. (2009). Comparison of a phased experimental approach and a single randomized clinical trial for developing multicomponent behavioral interventions. *Clinical Trials, 6*, 5–15.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755–764.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). (2013). *European drug report*. Publications Office of the European Union.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). (2014a). *Prevention profiles*. Retrieved January 23, 2014, from http://www.emcdda.europa.eu/prevention-profiles.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). (2014b). *About the best practice portal*. Retrieved January 23, 2014, from http://www.emcdda.europa.eu/best-practice/about.

Flay, B. R., Biglan, A., Boruch, R. F., et al. (2005). Standards for evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*, 151–175.

Gandhi, A. G., Murphy-Graham, E., Petrosino, A., Chrismer, S. S., & Weiss, C. H. (2007). The devil is in the details: Examining the evidence for "proven" school-based drug abuse prevention programs. *Evaluation Review, 31*, 43–74.

Griffin, K. W., Botvin, G. J., Scheier, L. M., Epstein, J. A., & Doyle, M. M. (2002). Personal competence skills, distress, and well-being as determinants of substance use in a predominantly minority urban adolescent sample. *Prevention Science, 3*(1), 23–33.

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., et al. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ, 336*, 924–926.

Hansen, W. B., Dusenbury, L., Bishop, D., & Derzon, J. H. (2007). Substance abuse prevention program content: Systematizing the classification of what programs target for change. *Health Education Research, 22*(3), 351–360. doi:10.1093/her/cyl091.

Heneghan, C. (2011). Considerable uncertainty remains in the evidence for primary prevention of cardiovascular disease [editorial]. The Cochrane Library 2011, 19 January. Retrieved August 31, 2011, from http://www.thecochranelibrary.com/details/editorial/983199/Considerable-uncertainty-remains-in-the-evidence-for-primary-prevention-of-cardi.html.

Hornik, R., Jacobsohn, L., Orwin, R., Piesse, A., & Kalton, G. (2008). Effects of the National Youth Anti-Drug Media Campaign on youths. *American Journal of Public Health, 98*, 2229–2236.

Kirby, D., Korpi, M., Barth, R. P., & Cagampang, H. H. (1997). The impact of the postponing sexual involvement curriculum among youths in California. *Family Planning Perspectives, 29*, 100–108.

Last, J. M. (2006). *A dictionary of Public Health*. New York: Oxford University Press.

Lipsky, M. S., & Sharp, L. K. (2011). From idea to market: The drug approval process. *The Journal of the American Board of Family Practice, 14*, 362–367.

Medical Research Council. (2000). *A framework for the development and evaluation of RCTs for complex interventions to improve health*. London: MRC.

Noar, S. M., & Zimmerman, R. S. (2005). Health behavior theory and cumulative knowledge regarding health behaviors: Are we moving in the right direction? *Health Education Research, 20*, 275–290.

Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2004). "Scared Straight" and other juvenile awareness programs for preventing juvenile delinquency. *Campbell Systematic Reviews*, *2*. doi: 10.4073/csr.2004.2

Pharmaceutical Research and Manufacturers of America (Pharma) (2007). Drug Discovery and Development: Understanding the R&D process. Retrieved July 21, 2011, from http://www.phrma.org/sites/default/files/159/rd_brochure_022307.pdf.

Pharmaceutical Research and Manufacturers of America (Pharma) (2011). Pharma—Drug Discovery and Development. Retrieved July 28, 2011, from http://www.phrma.org/research/drug-discovery-development.

Sackett, D. L. (2002). The arrogance of preventive medicine. *CMAJ, 167*, 363–364. http://www.ncbi.nlm.nih.gov/pubmed/12197692

Sallis, J. F., Mckenzie, T. L., Alcaraz, J. E., Kolody, B., Hovell, M. F., & Naderb, P. R. (1993). Project SPARK effects of physical education on adiposity in children. *Annals of New York Academy of Sciences, 699*, 127–136.

Sandler, I. N., Schoenfelder, E. N., Schoenfelder, S. A., & MacKinnon, D. P. (2011). Long-term impact of prevention programs to promote effective parenting: Lasting effects but uncertain processes. *Annual Review of Psychology, 62*, 299–329.

Sloboda, Z., Stephens, R. C., Stephens, P. C., Grey, S. F., Teasdale, B., Hawthorne, R. D., et al. (2009). The adolescent substance abuse prevention study: A randomized field trial of a universal substance abuse prevention program. *Drug and Alcohol Dependence, 102*, 1–10.

Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 Meta-analysis. *The Journal of Primary Prevention, 20*(4), 275–336.

Velleman, R. D., Templeton, L. J., & Copello, A. G. (2005). The role of the family in preventing and intervening with substance use and misuse: A comprehensive review of family interventions, with a focus on young people. *Drug and Alcohol Review, 24*(2), 93–109.

Valentine, J. C., Biglan, A., Boruch, R. F., Castro, F. G., Collins, L. M., Flay, B. R., et al. (2011). Replication in prevention science. *Prevention Science, 12*, 103–117.

Willis, T. A., Baker, E., & Botvin, G. J. (1989). Dimensions of assertiveness: Differential relationships to substance use in early adolescence. *Journal of Consulting and Clinical Psychology, 57*(4), 473–478.

Index

Z. Sloboda and H. Petras (eds.), *Defining Prevention Science*, Advances in Prevention Science, DOI 10.1007/978-1-4899-7424-2,