

ISSUES IN CLINICAL CHILD PSYCHOLOGY

Handbook of Research in Pediatric and Clinical Child Psychology

Practical Strategies and Methods

Edited by
Dennis Drotar

**Handbook of
Research in Pediatric
and Clinical Child
Psychology**

Practical Strategies and Methods

Issues in Clinical Child Psychology

*Series Editors: Michael C. Roberts, University of Kansas—Lawrence, Kansas
Lizette Peterson, University of Missouri—Columbia, Missouri*

BEHAVIORAL ASPECTS OF PEDIATRIC BURNS

Edited by Kenneth J. Tarnowski

CHILDREN AND DISASTERS

Edited by Conway F. Saylor

CONSULTING WITH PEDIATRICIANS: Psychological Perspectives

Dennis Drotar

HANDBOOK OF ADOLESCENT HEALTH RISK BEHAVIOR

Edited by Ralph J. DiClemente, William B. Hansen, and Lynn E. Ponton

HANDBOOK OF CHILD ABUSE RESEARCH AND TREATMENT

Edited by John R. Lutzker

HANDBOOK OF CHILD BEHAVIOR THERAPY

Edited by T. Steuart Watson and Frank M. Gresham

HANDBOOK OF CHILDREN'S COPING: Linking Theory and Intervention

Edited by Sharlene A. Wolchik and Irwin N. Sandler

HANDBOOK OF DEPRESSION IN CHILDREN AND ADOLESCENTS

Edited by William M. Reynolds and Hugh F. Johnston

HANDBOOK OF PSYCHOTHERAPIES WITH CHILDREN AND FAMILIES

Edited by Sandra W. Russ and Thomas H. Ollendick

HANDBOOK OF RESEARCH IN PEDIATRIC AND CLINICAL CHILD PSYCHOLOGY

Edited by Dennis Drotar

INTERNATIONAL HANDBOOK OF PHOBIC AND ANXIETY DISORDERS IN CHILDREN AND ADOLESCENTS

Edited by Thomas H. Ollendick, Neville J. King, and William Yule

MENTAL HEALTH INTERVENTIONS WITH PRESCHOOL CHILDREN

Robert D. Lyman and Toni L. Hembree-Kigin

SCHOOL CONSULTATION: Conceptual and Empirical Bases of Practice

William P. Erchul and Brian K. Martens

SUCCESSFUL PREVENTION PROGRAMS FOR CHILDREN

AND ADOLESCENTS

Joseph A. Durlak

A Continuation Order Plan is available for this series. A continuation order will bring delivery of each new volume immediately upon publication. Volumes are billed only upon actual shipment. For further information please contact the publisher.

Handbook of Research in Pediatric and Clinical Child Psychology Practical Strategies and Methods

Edited by

Dennis Drotar

*Rainbow Babies and Children's Hospital and
Case Western Reserve University
Cleveland, Ohio*

Springer Science+Business Media, LLC

Library of Congress Cataloging-in-Publication Data

Handbook of research in pediatric and clinical child psychology: practical strategies and methods/edited by Dennis Drotar.

p. ; cm. — (Issues in clinical child psychology)

Includes bibliographical references and index.

ISBN 978-1-4613-6869-4 ISBN 978-1-4615-4165-3 (eBook)

DOI 10.1007/978-1-4615-4165-3

I. Child psychiatry—Research—Handbooks, manuals, etc. 2. Sick children—Psychology—Research—Handbooks, manuals, etc. I. Drotar, Dennis. II. Series. [DNLM: I. Child Psychology. 2. Psychology, Clinical—Child. 3. Research—methods.]

WS 105 H2367 1999]

RJ499.3 .H367 1999

618.92'89—dc21

99-048699

ISBN 978-1-4613-6869-4

© 2000 Springer Science+Business Media New York

Originally published by Kluwer Academic / Plenum Publishers in 2000

Softcover reprint of the hardcover 1st edition 2000

All rights reserved

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher

Contributors

F. Daniel Armstrong, Department of Pediatrics, University of Miami, School of Medicine, Miami, Florida 33101

Leonard Bickman, Center for Mental Health Policy, Vanderbilt University, Nashville, Tennessee 37212

Ronald L. Blount, Department of Psychology, University of Georgia, Athens, Georgia 30602

Victoria L. Bunke, Department of Educational Psychology, University of Georgia, Athens, Georgia 30602

Colleen D. Carlson, Department of Psychology, University of Houston, Houston, Texas 77204-531

Sean D. Cleary, Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, Bronx, New York 10461

Julia M. Difilippo, Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44006-7123

Susan Douglas Kelley, Center for Mental Health Policy, Vanderbilt University, Nashville, Tennessee 37212

Dennis Drotar, Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038

Joseph A. Durlak, Department of Psychology, Loyola University—Chicago, Chicago, Illinois 60626

Debra Eisert, Oregon Health Sciences University, Child Development and Rehabilitation Center, Portland, Oregon 97207-0574

David J. Francis, Department of Psychology, University of Houston, Houston, Texas 77204-5341

Scott W. Henggeler, Family Services Research Center, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, South Carolina 29425

Nicholas S. Ialongo, Department of Mental Hygiene, Johns Hopkins University, Baltimore, Maryland 21205-1901

Anne E. Kazak, Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104-4939

Sheppard G. Kellam, Department of Mental Hygiene, Baltimore, Maryland 21205-1901

Gloria L. Krahn, Oregon Health Sciences University, Child Development and Rehabilitation Center, Portland, Oregon 97207-0574

Rachel Levi, Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123

Thomas R. Linscheid, Department of Psychology, Children's Hospital, Columbus, Ohio 43205

Anna T. Meadows, Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104-4939

Carol T. Nixon, Center for Mental Health Policy, Vanderbilt University, Nashville, Tennessee 37212

James C. Overholser, Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123

Tonya M. Palermo, Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038

Jeanne Poduska, Department of Mental Hygiene, Johns Hopkins University, Baltimore, Maryland 21205-1901

Alexandra L. Quittner, Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida 32610-0165

Jeff Randall, Family Services Research Center, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, South Carolina 29425

Kristin A. Riekert, Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123

Jane R. Robinson, Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038

Christopher Schatschneider, Department of Psychology, University of Houston, Houston, Texas 77204-5341

Anthony Spirito, Child and Family Psychiatry, Providence, Rhode Island 02903

Jane Timmons-Mitchell, Department of Psychiatry, University Hospitals, Cleveland, Ohio 44106

Natalie Walders, Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123

Laura L. Williams, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, South Carolina 29425

Kathleen Wells, Mandel School of Applied Social Sciences, Case Western Reserve University, Cleveland, Ohio 44106-7123

Thomas A. Wills, Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, Bronx, New York 10461

Jonathan F. Zaff, Department of Psychology, University of Georgia, Athens, Georgia 30602

Preface

The professional development of researchers is critical for the future development of the fields of pediatric and clinical child psychology. In order to conduct research in pediatric and clinical child psychology, researchers need to work with a wide range of populations and master an increasingly wide range of skills, many of which are either not formally taught or considered in sufficient depth in clinical training. Such skills include the development of resources for research by writing grants to government agencies and foundations; skills in preparing research for publications concerning original research, review articles, or case reports; scientific presentation skills; the ability to review and edit scientific manuscripts; and to implement and manage research in applied settings. Moreover, the increasing complexity of research in pediatric and clinical child psychology requires successful researchers in these fields to develop their expertise with a wide range of new specialized methodologies, data analytic methods, models of data analysis, and methods of assessment. Finally, to enhance the relevance of their research to practice, researchers in pediatric and clinical child psychology need to integrate their work with clinical service delivery programs that are based on empirical research.

The necessity to train researchers in pediatric and clinical child psychology in such multifaceted knowledge and skills places extraordinary burdens on professional training programs. Professional researchers in pediatric and child clinical psychology also are challenged to develop new knowledge and skills through continuing education and faculty development programs.

I have been involved in the professional training of scientist practitioners in pediatric and clinical psychology for the past 25 years. In trying to meet extraordinary challenges that are involved in training the next generation of researchers in pediatric and clinical psychology, I have become increasingly concerned about the need to provide information concerning the core skills and knowledge that will be needed to conduct research in these fields.

The present volume was developed from my experiences as a researcher and as a research mentor. These experiences consistently underscored the need to develop a handbook that would carefully consider the important, practical considerations in funding research, designing and implementing research, disseminating research, and integrating research and practice. Although these topics are important to the development of practicing researchers, they are not considered in available texts. In the course of implementing a research training program that has been

funded for the past 15 years by the National Institute of Mental Health and in which students in pediatric and clinical child psychology have participated, I often have found myself in the position of developing lectures and teaching materials to help students learn what I regarded core skills and knowledge as a researcher. This volume is partly a synthesis as well as extension of this work.

In organizing and developing this volume, I was guided by several questions: What had I learned about conducting research over the course of my career? What would I like and need to know about becoming a researcher if I were just starting my career? What skills are especially critical for researchers to develop the knowledge base in the field? The volume is organized around the following topic areas, each of which reflect critical areas of knowledge base and skills. These include: (1) using theoretical models and frameworks in data analysis, (2) designing research, (3) securing resources for research, (4) managing and implementing research in clinical settings, (5) disseminating research findings, (6) developing strategies to integrate research and practice, and (7) integrating research and policy.

In developing this volume, I chose a strategy that used a combination of chapters that were authored by myself and my students and that of expert researchers in clinical child and pediatric psychology whose work reflected key content areas as well as what I regarded as cutting-edge research. Each of the contributors to the volume was encouraged to prepare their chapter to emphasize a practical skill-based perspective; that is, to consider the problems that are faced by researchers in executing various tasks, to use illustrations from their work, and wherever possible to make specific recommendations for training researchers.

Readers should recognize that this volume is not intended as an exhaustive treatment of research methods in fields of clinical child and pediatric psychology but rather as a selective summary of key skill and knowledge areas. Given the breadth of the fields of pediatric and clinical child psychology, I faced a significant problem in choosing the sections to be covered and chapters to be considered in each of the sections. In making these decisions, I selected to focus on skill development and methodological issues that in my view were necessary to the professional development of researchers or represented new developments in methods or analytic approaches. I anticipate that subsequent revisions of this handbook will incorporate even greater richness of methods.

The inspiration for this book came from my experiences as a student and mentor. As a student I have been very fortunate to have worked with outstanding research mentors, Dick Lanyon, Ira Semler, and most especially, Don Routh. Now as a mentor, I have had the privilege to work with many talented students, among them, Laura Basili, Karen Berkoff, Beth Anne Bull, Barbara Boat, Erika Burgess, Marcy Bush, Carin Cunningham, Steve Evans, Yonit Hoffman, Carol Fitzpatrick, Carolyn Jivers, Carol Kucia, Rachel Levi, Amy Ludwig, Tonya Mizell-Palermo, Jack Nassau, Chantelle Nobile, Nancy Peterson, Charisse Peoples, Sean Phipps, Suzanne Powell, Kristin Riekert, Jane Robinson, Astrida Kaugars, Ron Saletsky, Lynn Singer, Lynne Sturm, and Natalie Walders.

ACKNOWLEDGMENTS. The hard work and dedication of Joyce Olson, who completed the word processing for much of this book, is gratefully acknowledged. My students, Erika Burgess, Rachel Levi, Kristin Riekert, Jane Robinson, Astrida Kaugars, and Natalie Walders, who reviewed various chapters, provided valuable feedback from the perspective of potential consumers of this book.

Contents

Part I. Applying and Testing Conceptual Frameworks and Theories in Data Analysis	1
Chapter 1. A Developmental Epidemiological Framework for Clinical Child and Pediatric Psychology Research	3
<i>Nicholas S. Ialongo, Sheppard G. Kellam, and Jeanne Poduska</i>	
Chapter 2. Testing Theoretical Models and Frameworks in Child Health Research	21
<i>Thomas A. Wills and Sean D. Cleary</i>	
Chapter 3. Introduction to Individual Growth Curve Analysis	51
<i>David J. Francis, Christopher Schatschneider, and Colleen D. Carlson</i>	
Part II. Designing Research and Applying Measures and Methods	75
Chapter 4. Understanding and Managing Sampling Issues in Research with Children	77
<i>Dennis Drotar and Kristin A. Riekert</i>	
Chapter 5. Strategies for Measurement and Psychological Assessment	97
<i>James C. Overholser, Anthony Spirito, and Julia M. DiFilippo</i>	
Chapter 6. Improving Assessment in Child Clinical and Pediatric Psychology: Establishing Links to Process and Functional Outcomes	119
<i>Alexandra L. Quittner</i>	
Chapter 7. Qualitative Methods in Clinical Psychology	145
<i>Gloria L. Krahn and Debra Eisert</i>	
Chapter 8. Understanding Cultural and Ethnic Influences in Research with Child Clinical and Pediatric Psychology Populations	165
<i>Natalie Walders and Dennis Drotar</i>	

Part III. Securing Resources for Research	189
Chapter 9. Writing Research Proposals for Foundations	191
<i>Dennis Drotar</i>	
Chapter 10. Preparing Grants to Secure Research Funding from Government Agencies	211
<i>Dennis Drotar</i>	
Part IV. Managing and Implementing Research with Pediatric and Child Clinical Populations	243
Chapter 11. Managing Research in Pediatric and Child Clinical Psychology	245
<i>Dennis Drotar</i>	
Chapter 12. Conducting Research with Children and Adolescents in Clinical and Applied Settings: Practical Lessons from the Field	261
<i>Dennis Drotar, Jane Timmons-Mitchell, Laura L. Williams, Tonya M. Palermo, Rachel Levi, Jane R. Robinson, Kristin A. Riekert, and Natalie Walders</i>	
Chapter 13. Multi-Institutional and Multidisciplinary Research Collaboration: Strategies and Lessons from Cooperative Trials	281
<i>F. Daniel Armstrong and Dennis Drotar</i>	
Chapter 14. Ethical Issues in Conducting Research with Pediatric and Clinical Child Populations in Applied Settings	305
<i>Dennis Drotar, James C. Overholser, Rachel Levi, Natalie Walders, Jane R. Robinson, Tonya M. Palermo, and Kristin A. Riekert</i>	
Part V. Disseminating Research Findings	327
Chapter 15. Presenting Scientific Data	329
<i>Dennis Drotar</i>	
Chapter 16. Writing Research Articles for Publication	347
<i>Dennis Drotar</i>	
Chapter 17. Writing Research Reviews	375
<i>Dennis Drotar</i>	
Chapter 18. How to Evaluate a Meta-Analysis	395
<i>Joseph A. Durlak</i>	

CONTENTS	xiii
Chapter 19. Reviewing and Editing Manuscripts for Scientific Journals	409
<i>Dennis Drotar</i>	
Part VI. Developing Strategies to Integrate Research and Practice	425
Chapter 20. Case Studies and Case Series	429
<i>Thomas R. Linscheid</i>	
Chapter 21. Conducting Randomized Treatment Studies in Real-World Settings	447
<i>Scott W. Henggeler and Jeff Randall</i>	
Chapter 22. Evaluating Mental Health Services for Children and Adolescents	463
<i>Susan M. Douglas Kelley, Carol T. Nixon, and Leonard Bickman</i>	
Chapter 23. The Integration of Basic Research, Treatment Research, and Clinical Practice in Pediatric Psychology	491
<i>Ronald L. Blount, Victoria L. Bunke, and Jonathan F. Zaff</i>	
Chapter 24. Integrating Psychosocial Research and Practice in a Pediatric Hospital	511
<i>Anne E. Kazak and Anna T. Meadows</i>	
Part VII. Integrating Research and Policy	529
Chapter 25. Use of Research in Mental Health Practice and Policy	531
<i>Kathleen Wells</i>	
Index	545

**Handbook of
Research in Pediatric
and Clinical Child
Psychology**
Practical Strategies and Methods

I

Applying and Testing Conceptual Frameworks and Theories in Data Analysis

Conceptual frameworks and theories have several important advantages to develop and frame research questions in most areas of science. The fields of pediatric and clinical child psychology are no exception. Theories and framework can serve to clarify investigators' thinking about difficult research problems. Moreover, theories and frameworks can help researchers organize and integrate information from disparate studies into a coherent form. Finally, theories and frameworks facilitate the development of testable hypotheses that are critical to successful research protocols. Yet, despite these advantages, theories and frameworks have not been used to full advantage by researchers in pediatric and clinical child psychology.

OBSTACLES TO THE USE OF THEORY

What accounts for these practices? While one cannot know for sure without having access to hearts and minds of researchers, several hypotheses can be proposed. One of these is that pediatric and clinical child psychology are relatively new fields. Consequently, there has not been sufficient time to develop theories based on programmatic research. The applied nature of pediatric or clinical child populations and practice also may limit the use of theory. Pragmatic, nose-to-the-grind researchers who are working on applied problems may not always recognize or appreciate the value of a theory. Some researchers have a stronger interest in applying methods of assessment and intervention than they have in understanding the processes by which assessment and interventions work. Finally, researchers may not utilize theories or frameworks because they have not been taught how they can be used to guide research. Another explanation for the relative underutilization of theory may relate to investigators' concerns that their cherished theories may not be confirmed if it is put to a solid test.

To illustrate how theory can be used in research, the first section of the book considers how theoretical frameworks and models can be used to guide the

development of programmatic research and an informed approach to data analysis. In Chapter 1, Ialongo, Kellam, and Poduska describe the principles and theoretical perspectives of a comprehensive developmental epidemiological framework that has formed the foundation for the Johns Hopkins University Prevention Intervention Research Center's research programs concerning children's mental health. These researchers describe the utility of a community epidemiology perspective in developing preventive interventions that address the impact of the ecology of the school classroom and the impact of neighborhood social ecology on children's mental health. The importance of a public health theoretical perspective in research is also illustrated in the development of collaborative preventive intervention trials. The utility of a developmental perspective included in this model is shown by the tests of the developmental functions of the target antecedent behaviors of academic underachievement in children. These authors conclude with recommendations for training psychologists in public health and epidemiological perspectives.

In Chapter 2, Wills and Cleary present a useful description of testing theoretical models and frameworks in child health research. Based on illustrations from a research program concerning the origins and consequences of substance use in adolescents and families, Wills and Cleary describe alternative theoretical models, for example, social influence, stress and coping, and types of data-analytic models, including, direct effects versus mediation models, which are available to test these theoretical models. Using examples from research on socioeconomic disadvantage as a moderator for adolescents' adjustment, childhood chronic illness, and maternal mental health, these authors take the researcher step by step in constructing and testing a theoretical model using structural equation modeling. Throughout the chapter, Wills and Cleary give the reader practical suggestions concerning analytic strategies for testing models.

Many of the important research questions in clinical child and pediatric psychology involve theoretical models and analytic approaches to assess change, such as assessment of the effectiveness of intervention in long-term adaptation to chronic illness and so forth. Individual growth models for measuring change and examining correlates of change provide an important alternative to limitations of traditional approaches to assess change. In Chapter 3, using the illustration of a study of prediction of growth in children's reading achievement skills across a nine-year time span (grades 1 through 9), Francis, Schatschneider, and Carlson provide an introduction to growth curve modeling for researchers. This chapter demonstrates the use of this important analytic tool to describe and evaluate change in individual children.

1

A Developmental Epidemiological Framework for Clinical Child and Pediatric Psychology Research

**NICHOLAS S. IALONGO, SHEPPARD G. KELLAM,
and JEANNE PODUSKA**

INTRODUCTION

In this chapter we elaborate on the principles and theoretical perspectives that have formed the foundation for the Johns Hopkins University Prevention Intervention Research Center's (JHU-PIRC) work over the last decade. During this time, we fielded two sets of classroom-based, universal preventive intervention trials in 28 Baltimore City schools with three cohorts of first graders. We also fielded a universal, or primary, preventive intervention trial, along with a nested mental health services intervention, in three Baltimore Head Start centers. In each trial, the focus was on the early risk behaviors of poor academic readiness and aggressive and shy behavior and their distal correlates of antisocial behavior, substance abuse, and anxious and depressive symptoms and disorders. We believe the principles and perspectives that have guided our preventive and services interventions work are equally applicable to researchers in the child clinical and pediatric psychology arenas. As will become apparent, we advocate an interdisciplinary approach to mental health research, which requires an integration of the life course developmental, public health, community epidemiological, and experimental trials perspectives. We provide the reader with examples from our own research as well as

NICHOLAS S. IALONGO, SHEPPARD G. KELLAM, and JEANNE PODUSKA • Department of Mental Hygiene, Johns Hopkins University, Baltimore, Maryland 21205-1901.

that of others that we hope will highlight the strengths of our developmental epidemiological framework and its applicability to the challenges faced by child clinical and pediatric psychology researchers. Among the topics addressed are: (1) the need for grounding our clinical and basic research in developmental theory (Coie et al., 1993; Kellam & Rebok, 1992; Reid, 1993; Sandler et al., 1992); (2) the use of intervention trials as tests of our etiologic models of development and psychopathology (Coie et al., 1993; Kellam & Rebok, 1992; Reid, 1993; Sandler et al., 1992); (3) the benefits of a public health perspective with respect to the dissemination and acceptance of interventions found to be efficacious in research settings (Kellam & Rebok, 1992); (4) the pitfalls associated with the use of clinic samples and samples of convenience in clinical research and research on development and psychopathology (Greenley & Mechanic, 1976; Greenley, Mechanic, & Clearly, 1987); and (5) economically and logically feasible methods for defining and selecting research populations in accord with epidemiological principles (Gordis, 1996).

THE LIFE COURSE-SOCIAL FIELDS FRAMEWORK

The JHU-PIRC's decade-long approach to the prevention and treatment of mental disorders has been to target the early antecedent risk behaviors of these disorders through a program of distinct but complementary preventive and mental health services interventions. The choice of the proximal targets for the center's interventions and their timing and assessment have been guided by our life course-social fields framework (see Fig. 1) (Kellam & Rebok, 1992; Mrazek & Haggerty, 1994), which is based on an integration of four perspectives: life course development, community epidemiology, preventive intervention trials, and public health. The JHU-PIRC's adoption of an interdisciplinary perspective is consistent with

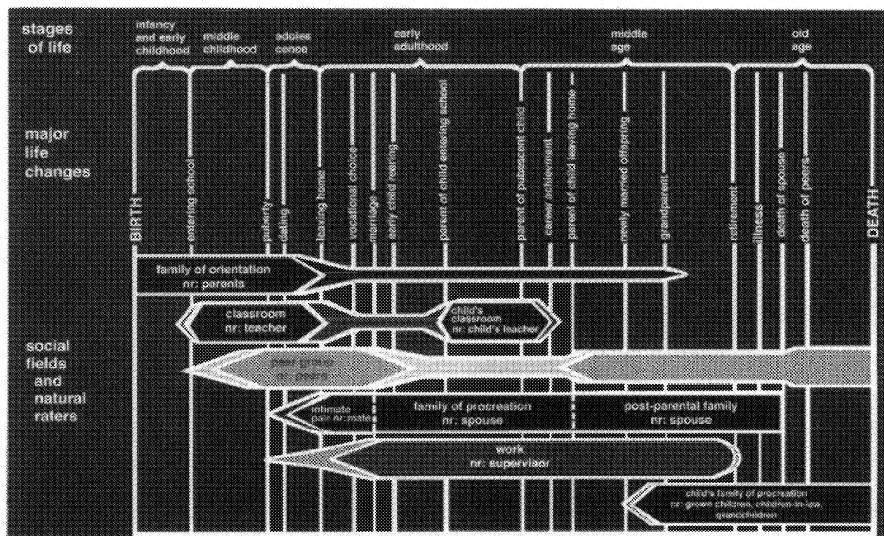


Figure 1. Life course social field framework (Kellum et al., 1975).

recent advances in mental health research that have caused investigators to question to whom their research findings pertain and the role of context in shaping human development. Indeed, current sampling procedures often leave uncertain the populations of individuals or families to whom the research findings can be generalized. That is, the sample is not representative of a defined population. The fact that the frequency and distribution of the causal processes that put children and their families at increased risk may vary across social contexts such as neighborhoods or communities serves to highlight the need for defining the population under study. In our view of prevention and mental health services research, the need for defining the population under study is particularly important, since the causal model provides the specific targets for our preventive and mental health services interventions. For example, a preventive intervention aimed at reducing the risk of antisocial behavior is likely to be of little value to a community if the prevalence of the targeted risk factors (e.g., exposure to deviant peers) or processes (e.g., inept parenting practices) within that community is quite low. An alternative causal model and preventive intervention may be a better fit for that community. Finally, the JHU-PIRC's adoption of an interdisciplinary perspective stems from the fact that the causal models offered by clinical researchers are frequently limited by the absence of relevant aspects of the environmental context, as we will elaborate upon below.

Life-Course Development

As indicated above, the life course–social field framework reflects a life course development orientation. Life course development focuses on the mapping of developmental paths including antecedents, mediators, and moderators of developmental processes and consequences (Kellam, Branch, Agrawal, & Ensminger, 1975; Kellam & Ensminger, 1980; Kellam & Rebok, 1992). Research on developmental paths that includes searching for antecedents and elements that enhance or inhibit developmentally appropriate outcomes are viewed in the context of variation in individuals within defined populations. The purpose of carrying out research on developmental paths is to uncover aspects of developmental models that may be important in the developmental and etiologic outcomes and that are amenable to intervention trials.

Central to life course–social field theory is the concept that individuals face specific social task demands in various social fields across the major periods of the life span (Kellam & Rebok, 1992). The social task demands that the individual confronts are defined by individuals in each social field, whom we have termed the natural raters. The natural rater not only defines the tasks but rates the individual's performance in that social field. Parents function as natural raters in the family, peers in the peer group, and teachers in the classroom (Kellam, 1990; Kellam et al., 1975; Kellam & Ensminger, 1980). This interactive process of demand and response is termed *social adaptation* and the judgments of the individual's performance by the natural raters *social adaptational status* (SAS) (Kellam et al., 1975). In line with the organizational approach to development (Cicchetti & Schneider-Rosen, 1984), normal development is viewed within the life course–social fields framework as marked by the integration of earlier competencies into later modes of function, with the earlier competencies remaining accessible, ready to be activated and utilized during times of stress, crisis, novelty, and creativity. It follows then that early successful social adaptation in the face of prominent developmental challenges

tends to promote later adaptation as the individual traverses the life course and encounters new and different social task demands across the main social fields (Cicchetti & Schneider-Rosen, 1984). This key developmental principle along with a growing empirical literature forms the basis for the JHU-PIRC's focus on successful adaptation to first grade as a means of improving social adaptational status over the life course.

In contrast to SAS, psychological well-being (PWB) in the life course–social fields framework refers to the individual's internal state, as reflected in anxious and depressive symptoms and mood disorders. We hypothesize that PWB and SAS are intimately related, such that PWB is in large part determined by the degree to which the individual is successful in meeting the demands of his or her natural raters. Our conceptualization of the link between SAS and PWB is grounded in the basic principles of the social learning theories of depression. The more successful an individual is in meeting the demands of his or her natural raters, the more likely he or she will be reinforced for his or her successes. Alternatively, failure to meet the demands of the natural raters will be associated with reductions in reinforcement and increased punishment, which then may lead to decrements in psychological well-being. This hypothesized link between SAS and PWB provides additional rationale for targeting early SAS as a means of not only improving subsequent SAS but of reducing the risk for decrements in PWB and early and sustained substance use as well.

COMMUNITY EPIDEMIOLOGY

The community epidemiological perspective is also represented in our life course–social field framework. Community epidemiology is concerned with the nonrandom distribution of a health problem, or related factor, in a fairly small population in the context of its environment, such as a neighborhood, school, or classroom. Community epidemiology informs us as to variation in developmental paths, including the roles of antecedents, mediators, and moderators, as they vary in frequency and function within and across different subgroups and contexts of a defined population. Traditionally, the phrase host–agent–environment is part of the epidemiological lexicon (Morris, 1975). It refers to a way of conceptualizing cause or etiology as involving vulnerability in the person (the host), conditions in the environment as producing illness, and a causal process of interaction (the agent) between the individual and environmental risk conditions. Thus, the integration of life course development with community epidemiology allows the study of variation in developmental antecedents and paths in a defined population in defined ecological contexts.

Defining the Ecological Context

From a community epidemiological perspective, a neighborhood can be defined in terms of its geographic boundaries and its sociodemographic characteristics, particularly those social indicators that may be relevant to mental health. The health demographic profile (Goldsmith et al., 1984) developed by the National Institute of Mental Health's (NIMH) Biometry and Epidemiology branch, using census data, provides a means of rapidly characterizing neighborhoods with re-

spect to small-area social indicators that have been found to be related to the incidence and prevalence of mental disorder. In fact, the original health demographic profile was developed to determine the need for community mental health centers in neighborhoods and communities across the country based on small-area social indicators. That is, the small-area social indicators were used to predict the rate of mental disorder in the community and to estimate the need for mental health services based on that rate.

The Advantages of Defining the Ecological Context

With the use of samples of convenience or clinic samples, as is often the case in child clinical and pediatric psychology research, subjects tend to be viewed in isolation from their socioenvironmental characteristics. For example, the characteristics of a child's classroom, peer group, family, or neighborhood cannot be included as precisely as needed to understand variation in intervention impact or variation in the developmental course. Our community epidemiological perspective is in keeping with Bronfenbrenner's (1979) admonition to consider the determinants of human development arising from the broader environment in which children and families are embedded.

Effects of the Classroom Social Ecology

Kellam et al. (1998) found that the level and duration of response to school-based, preventive interventions may vary as a function of the characteristics of the child's classmates and of the classroom and school. More specifically, they found the risk of being rated as highly aggressive in middle school for boys varied as a function of the level of aggression in the first grade classroom, after controlling for the youth's level of aggression in first grade. That is, controlling for boys' level of aggression in first grade, boys in first grade classrooms rated as highly aggressive were at four times greater risk of being rated as aggressive in sixth grade than boys who were in low-aggressive first grade classrooms. Similarly, in our most recent elementary school-based preventive intervention trial targeting aggressive behavior, the largest effect found was the intervention condition by school interaction, with schools with the highest levels of behavior problems prior to the intervention being the least likely to show an intervention effect (Ialongo et al., *in press*). One possible explanation for this result is that if the prevailing behavior in the classroom is disruptive and aggressive, aggression may be tolerated to a greater degree by both teachers and peers and therefore may go unpunished. As a result, the likelihood of future acts of child aggression may be increased. Relatedly, high rates of disruptive and aggressive behavior may cause teachers to react in a more irritable and coercive fashion with their students. When this happens, the students may reciprocate by increasing the frequency and/or amplitude of their disruptive and coercive behavior. As a further example of the potential impact of classroom-school characteristics on intervention outcomes, large class sizes may serve to reduce teachers' capacity to adequately and consistently monitor and discipline each of their students. Moreover, either of the above—high rates of disruptive behavior and/or large class sizes—may result in teachers spending less time on rehabilitative work with students who are falling behind academically and/or who are aggressive.

Effects of the Neighborhood Social Ecology

A community epidemiological perspective also leads us to examine factors operating at the level of neighborhood that may influence the risk for a mental disorder or the effects of preventive or services interventions (Brook, Nomura, & Cohen, 1989). In the case of the former, the risk of depression for African-American adults in the NIMH epidemiologic catchment area (ECA) studies varied as a function of the racial composition of the neighborhood they lived in (Tweed et al., 1990). If you were African American and lived in a majority African-American neighborhood, you were at lower risk for depression than if you were in a majority “white” neighborhood. Without careful definition of the population and its social context, such phenomena may not have been discovered.

As to neighborhood influences on the effects of preventive and services interventions, let us take, for example, an intervention aimed at preventing substance use. Whether such an intervention is successful may vary with the availability of substances in the neighborhood. The greater the availability, the greater likelihood of use. Consequently, we may see poorer intervention response in neighborhoods with high availability (Johnston, O’Malley, & Bachman, 1995; National Household Survey of Drug Abuse, 1995). In addition, the individual youth’s attitudes and beliefs about substance use may be shaped by the prevailing attitudes and beliefs at the level of the neighborhood. In a neighborhood where the prevailing attitudes are accepting of substance use, the individual’s attitudes may become more accepting as well. Importantly, Johnston et al. (1995) found that as disapproval and perceptions of harm of marijuana use have decreased since 1992, use of marijuana has increased. Relatedly, Crum, Lillie-Blanton, and Anthony (1996) report Baltimore City youths living in neighborhoods in the highest tertile of crime and drug use were 3.8 times more likely to have been offered cocaine than youths in the lowest tertile.

In terms of neighborhood influences on preventive interventions aimed at educational outcomes, children in neighborhoods characterized by high levels of unemployment may perceive that regardless of their academic efforts and successes, high-paying jobs may be unattainable once they enter the workforce. Consequently, they may be less likely to demonstrate sustained academic effort and more likely to drop out of school. An additional factor operating at the level of neighborhood that may influence intervention response is the availability of formal support systems, such as affordable, quality child care services, and well-supervised after-school programs that provide children with opportunities to engage in appropriate educational, recreational, and social activities. Finally, the availability of child and family mental health services may serve to influence intervention response through direct facilitation of adaption to normative developmental demands and/or by facilitating children’s coping with failure to meet task demands, either through psychosocial or pharmacological means.

The Community Epidemiological Approach versus the Use of Weighted National Samples

Community epidemiology (as distinguished from the use of weighted national samples) is well suited for analytic and explanatory goals of preventive and mental

health services research. More specifically, utilizing community epidemiological principles and methods, we can hold constant the macro-characteristics of a population, for example, an urban neighborhood. We can then examine diverging developmental paths in the context of variation in small social fields such as family, classroom, and classmate-peer group within that neighborhood. In contrast to the community epidemiological approach, the use of weighted national probability samples leaves us with too few cases in any one ecological context to study the effects of that context on development.

Community Epidemiology and Sampling

The community epidemiological perspective offers a number of advantages, particularly with respect to sampling. Volunteer samples, or samples drawn from clinics, come from unknown total populations. Such samples typically entail selection bias, since those families who volunteer or who seek help may be different in important aspects from families with similar problems who do not (Greenley & Mechanic, 1976; Greenley et al., 1987; Kellam, Branch, Brown, & Russell, 1981). Those who seek help from the church may be quite different from those who seek help from the clinic. Subjects in volunteer or clinic samples differ from the general population by the very fact that they seek help (Kellam et al., 1981). Relying on volunteer subjects in prevention or services intervention trials sought through newspaper or poster advertising has similar problems. Those who respond may not be representative of those who do not. Of note, the work of Leaf, Alegria, and Cohen (1996) and others (Offord, Boyle, & Szatmari, 1987; Zaner, Pawelkiewicz, DeFrancesco, & Adnopo, 1992) suggest that less than 25% of children and adolescents in need of mental health services receive such services. Thus, the children who do come to mental health specialty clinics represent only a fraction of the population with mental health problems.

Among the most critical of the potential biases associated with the use of clinic samples is the tendency for only the most socially impaired of children to be referred to and seen in clinics (Berkson, 1946; Caron & Rutter, 1991). Indeed, Caron and Rutter (1991) demonstrate that the prevalence of psychiatric disorders and their comorbidity are vastly overestimated when clinic-based samples are used. In avoiding this bias, epidemiologically defined, community samples ensure that generalizations to known populations can be drawn and the degree of social and cognitive impairment associated with psychiatric symptoms, syndromes, or disorders, along with their incidence, prevalence rates, and comorbidity, can be validly inferred (Kellam, 1990). In the absence of such data, informed decisions with regard to the allocation of the limited treatment and preventive intervention services available are difficult to make (Kellam, 1990; National Institute of Mental Health, 1991). Thus, mental health service needs may go unmet.

An Understanding of Who Participates and Who Does Not

An additional advantage of the community epidemiological perspective is that few samples are likely to be complete in the sense of all of the targeted population being constantly and continuously available. Prevention or treatment research with children is particularly difficult given the mobility of families. Here, a com-

munity epidemiological orientation providing information about the total population offers an understanding of who participates compared to those who do not.

The above represent some of the most important reasons why our prevention and mental health services research is oriented to the community epidemiological perspective. Community epidemiology provides the methodology for obtaining population rates and distributions of antecedents and outcomes and tools and concepts for integrating disciplines into a broader, more ecological perspective for clinical and pediatric research.

THE PUBLIC HEALTH PERSPECTIVE

In addition to life course development and community epidemiology, our life course–social fields framework features a public health perspective. An important advantage of our public health perspective is that the diffusion of effective programs is facilitated by partnerships fostered with the major institutions charged with the public's health, education, and welfare. From the beginning, the JHU-PIRC preventive and mental health services efforts have been developed in conjunction with personnel from the institutions expected to implement them and integrated into the ongoing activities of those institutions. Intentionally, we seek to ensure that once the research funds are no longer available, the institution retains a trained cadre of intervenors with the materials and protocols necessary to sustain effective programs. In line with our public health perspective, one way we sought to insure the dissemination and acceptability of interventions was to enter into a partnership with the existing, public institution mandated by the city, county, and state to socialize and educate Baltimore City's children—the Baltimore City public school system (BCPS). Indeed, the JHU-PIRC's partnership with the BCPS has evolved over the last 12 years to the point that members of the JHU-PIRC team were integrated into the BCPS's curriculum, parent involvement, special education, and mental health services planning committees. Consequently, the first grade interventions implemented in the latest JHU-PIRC school-based preventive intervention field trials represented not only what the BCPS thought was affordable and feasible, but the directions in which the BCPS was going in terms of new initiatives in the areas of curriculum, parent involvement efforts, and mental health services. Each element of the interventions reflected the thinking of the BCPS superintendent, his administrators, principals, and teachers. In addition, each element of the interventions was piloted and feedback solicited not only from principals, teachers, and school social workers/psychologists, but from parents and children as well. Moreover, rather than importing "experts" to provide preventive services, existing school staff—principals, teachers, school psychologists, and social workers—collaborated on the development, implementation, and evaluation of the center's preventive intervention trials over the last 10 years. Thus, this public health approach allows the JHU-PIRC to confirm the applicability of findings from laboratory and microanalytic studies to population settings and to ensure that the public benefits of large-scale interventions outweigh public costs (Kellam, Rebok, Mayer, Ialongo, & Kalodner, 1994a; Kellam, Rebok, Ialongo, & Mayer, 1994b). Strong collaborative partnerships are also necessary for population-based intervention trials requiring random assignment of teachers, children, and families to intervention and control conditions.

Pentz (1993) and others (Wagenaar, Murray, Wolfson, Forster, & Finnegan, 1994; Perry et al., 1996) offer examples of large-scale community preventive intervention efforts involving strong collaborative partnerships with local, county, and state-wide institutions. Minkler (1990) provides a theoretical model for developing collaborative relationships with the institutions integral to the public's mental health, along with the pragmatics of achieving a successful collaboration. Jason (1982) has described similar partnership strategies for developing community and institutional support. Such partnerships require considerable time to build, along with a mutual sense of trust and shared interests (Kellam et al., 1975). Our initial intervention efforts in the BCPS required a 30-month development period. Such prolonged start-up times may serve to deter most clinical and/or pediatric researchers. But once such collaborations are developed and prove productive to the institutions involved, start-up times for future research efforts are often dramatically reduced.

A Schema for Organizing Preventive and Treatment Services

Consistent with our public health perspective, the schema for organizing our preventive and services efforts has involved four levels of interventions nested in a public health–human services system. At the first level, universal interventions (Mrazek & Haggerty, 1994) are applied to all children and families. At the second level, selective interventions back up the universals for those children who require more help. Indicated interventions are at the third level and back up the first two levels. At the fourth and final level are treatment services for those children and families who fail to benefit from the preventive interventions and who have mental disorders. The first, or universal level, addresses the socialization structure and processes by which public institutions such as schools foster child social, cognitive, emotional, and behavioral development. The settings for universal interventions include the institutions of family, Head Start classroom, the public school system, and the myriad agencies providing mental health and social services to children and their families. Response to the universal interventions serves as a means of reliably identifying children and families in need of additional and/or more intensive intervention at either the selective, indicated, or treatment level. The selective level is intimately related to the first, or universal, level, and entails more specialized professional care. The third, or indicated level, is still more specialized and typically more expensive and requires a more remedial focus rather than strengthening the socialization structures in institutions such as Head Start, grade school, and/or the family. The fourth level, treatment services, typically involves the provision of highly specialized habilitative or rehabilitative care, within traditional mental health treatment settings.

Our current Head Start intervention trials embodies this nested approach to preventive and mental health services. We have developed and are now evaluating a universal, Head Start, parenting preventive intervention, which utilizes paraprofessionals as leaders and targets those parenting practices associated with the early antecedent risk behaviors of aggressive and shy-withdrawn behavior and their later correlates: antisocial, affective, and drug and alcohol disorders. In addition, we have developed and are now evaluating a system for assessment and identification of mental health service needs and linkage to mental health services for Head Start children and families. This system of identification and linkages

builds on the integration of a mental health specialist into a Head Start interdisciplinary team. This team is responsible for seeing that the mental health needs of Head Start families and children are met along with their social and physical health needs. Through the Head Start mental health specialist and the Head Start interdisciplinary team, we have developed a network of linkages to mental health and substance abuse treatment providers in Baltimore city. We also have developed a set of first-stage, or screening, assessment tools that are currently being used by Head Start teachers and family service coordinators to identify Head Start children and families who may be in need of mental health services. These children and families identified to be in need based on these first stage measures are then referred to the Head Start mental health specialist for a more comprehensive assessment and the development of a treatment plan. In the event that the plan calls for linkage to mental health service providers in the community, the mental health specialist serves as the family's advocate and liaison to those agencies and remains in a case manager role throughout the process. In addition to assessment and linkage to services, the role of the mental health specialist includes: (1) training and supervising Head Start family service coordinators and teachers in identification of children and families in need of mental health services; (2) consulting with and training Head Start teachers on the use of effective classroom behavior management strategies; (3) training and supervising the paraprofessionals who lead the universal parenting intervention; (4) providing on-site, time-limited mental health services to Head Start families and children; (5) developing and maintaining a computer database of available mental health services in the community; and (6) the establishment of contractual agreements with the local mental and substance abuse treatments providers for on-site services whose costs are covered through third-party billing.

As an example of how of our nested approach to prevention and treatment works, a parent attending one of the universally offered Head Start workshops acknowledges great difficulty in applying the principles discussed in the workshops. The Head Start teacher concomitantly reported, as part of our first-stage assessments, that the child's behavior in the classroom had worsened since the year began. The workshop leader and Head Start teacher then consulted with the Head Start mental health specialist, who recommended and performed a comprehensive assessment of the parent, child, and family. The assessment revealed that the parent was suffering from depression and substance abuse. The Head Start mental health specialist subsequently assisted the parent in obtaining mental health and drug treatment services. In addition, the mental health specialist developed with the Head Start teacher an individualized classroom behavior change program for the child, which complemented the ongoing universal classroom-behavior management intervention. The mental specialist also saw the child individually and worked on helping the child cope with his mother's illness.

PREVENTIVE INTERVENTION TRIALS AS EXPERIMENTAL TESTS OF DEVELOPMENT MODELS

The fourth perspective that is reflected in our life course-social fields framework is preventive intervention trials as experimental tests of developmental models (Kellam & Rebok, 1992). Our preventive trials in the Baltimore public

schools are directed at early behavioral responses of children to school requirements that have been demonstrated in prior research to be risk factors for later antisocial behavior, substance abuse, and depression. The results of our preventive trials provide tests of the developmental significance of early risk behaviors or conditions as well as specific intervention impact. Our goal has been to determine whether the risk of the outcome decreases if the antecedent predictor changes. An affirmative answer informs our understanding of the etiologic significance of the predictor, whether it is a behavioral response or a family process. Thus, intervention trials can be directed at changing the target antecedent to test its malleability as well as the consequences of its change on developmental outcomes. In addition, such trials not only allow the exploration of the variation in developmental paths, but also the differences among the responders and the nonresponders to the preventive intervention. Knowledge of these can inform the next stage of intervention programming. In essence, then, the intervention trial has the dual purpose of experimentally testing the developmental functions of the target antecedent behaviors and of developing effective preventive programs (Kellam & Rebok, 1992; Kellam & Werthamer-Larsson, 1986). This experimental approach (Morris, 1975) integrates population-based intervention research with research on etiology, development, and vulnerability.

As an example of a preventive intervention trial informing research on etiology and vulnerability, Kellam et al. (1994a), based on life course–social fields theory, hypothesized that first graders' psychological well-being was integrally tied to their success in meeting the academic demands of the first grade classroom. A school-based preventive trial aimed at improving academic achievement in first grade was fielded, and both academic achievement and psychological well-being were assessed at pre- and posttest in the fall and spring of first grade, respectively. Schools were randomly assigned to intervention and control conditions. At posttest in the spring of first grade, Kellam and colleagues examined whether the beneficial effect of the intervention on psychological well-being during first grade was mediated by the gains in academic achievement found for the intervention participants versus the controls. The results supported the hypothesis that children's psychological well-being was in fact tied to their academic performance in first grade, although not exclusively, which is consistent with the view that psychological well-being is a multidetermined phenomena.

Developmental Epidemiology and the Efficiency of Multistage Sampling

Research on developmental antecedents, paths, and outcomes from our life course–social fields perspective requires population-based measurement designs that control for selection bias in ways that samples derived from clinics or from voluntary samples cannot. In turn, developmental epidemiologically based research requires ecologically valid and economical measures that provide important information about the characteristics of the individuals in the population and the antecedents, mediators, and moderators of developmental processes and consequences. These measures should also inform us as to the extent to which developmental outcomes and their putative mediators and moderators vary in frequency and function within and across different subgroups of a defined population in the context of a community.

An important strategy for maximizing efficiency in epidemiological research is

the multistage sampling strategy (Anthony, 1990). It conserves resources by using efficient assessments of large, population-based, probabilistic samples and more expensive and burdensome assessments on subsamples selected by reason of supposed high risk (Anthony, 1990). Through multistage, representative sampling, the data from our first-stage, or population-based, measures can be used to draw smaller samples for studies requiring more frequent and comprehensive measurement and close laboratory control. Population estimates for the more intensive measures can be made through use of sample selection weights (Anthony, 1990).

First-stage, or population-based, measures have served three key functions in our current and previous developmental epidemiological field trials. First, they have been relied on to provide measures of intervention effects and outcomes. Consequently, they needed to be reliable and sensitive measures of change that can be briefly and economically administered to the entire population of interest. Second, they have been used to identify individuals from the population in need of selected or indicated preventive interventions or treatment. As noted above, we advocate a nested approach to preventive and mental health services interventions, with universal preventive interventions serving as routing mechanisms to selective and/or indicated interventions. That is, our first-stage measures are used to measure response to the universal interventions, and as such serve to identify the children in need of more intensive mental health services.

A third function of our first-stage measures is that they serve to provide the needed bridge for linking developmental epidemiology to studies based on more frequent or precise observations on smaller samples. That is, research on total cohorts within specified populations can be related through multistage representative sampling to microanalytic studies on selected smaller populations requiring more frequent measurement and close laboratory control. For example, a smaller, stratified probability sample was drawn from our first JHU-PIRC cohort to study attention-concentration processes in young children more comprehensively. The sample was selected using relevant first-stage measures that were administered to the entire cohort briefly and economically. This second-stage sample was drawn to represent the strata from which it was derived as well as the total population. The first-stage measures revealed an intimate relation between teacher-rated concentration problems and aggressive as well as shy behaviors, classroom achievement, and depression. The representative smaller sample studied periodically with more precise measures of attention confirmed the associations and specified the kinds of attention deficits involved (Mirsky et al., 1991). These results then led to the inclusion of attention deficits in our analyses of impact of the JHU-PIRC's preventive trials and to the development of new trials for children at risk for attention disorders.

As Kellam (1990) has noted, this same approach could be used to identify cases of interest for child clinical and pediatric psychology researchers. For instance, Ialongo, Reider, and Kellam (to be published) were interested in studying the prognostic power of young children's self-reports of depressive and anxious symptoms. Detailed psychiatric assessments were carried out in the context of a multi-stage sampling design. At the first stage, a relatively economical yet sensitive screening instrument was used to identify cases for more precise and comprehensive assessment at a second stage. A random stratified sampling procedure was then employed at the second stage to select representative samples of children from the

entire distribution of scores on the first stage measure of depressive symptoms. Bird, Gould, Yager, Staghezza, and Canino (1989) and others (e.g., Offord, Boyle, & Racine, 1989) provide examples of the use of multistage designs to estimate the prevalence of psychiatric disorders in children and adolescents. Bird et al. (1989) and Offord et al. (1989) used the Child Behavior Checklist (CBCL) (Achenbach, 1991) at the first-stage, or population level. The administration of the CBCL was followed by clinical interviews at the second stage for those children identified as cases by the CBCL and a random-stratified sample of controls (Bird et al., 1989; Offord et al., 1989).

Anthony (1990) describes how a multistage sampling strategy can be incorporated into a longitudinal, prospective design. At each longitudinal assessment new cases are identified through the use of the first-stage measure(s) administered at the population level. For each new case, matched controls are drawn from the remaining noncase population for more intensive study, along with the previously identified cases and their second-stage controls. This link between population-based epidemiological research and the smaller samples required for microanalytic study shows great promise for future research in pediatric psychology and the child clinical areas along with the next stage of prevention research.

Each of the designs discussed above are variants of the case-control design, which provides an efficient means of studying relatively rare disorders or phenomena (Lilienfeld & Stolley, 1994; Schlesselman, 1981). In the case-control design, a suitable number of matched controls are drawn for each case. The controls may be drawn from the same classroom or school or neighborhood, depending on the investigator's hypotheses with respect to the level of the ecological context that will contribute the most to understanding the phenomena of interest. Snow's (1849) work on the outbreak of cholera in 19th-century London aptly illustrates the advantage of a case-control design. Snow discovered while comparing cases of cholera to cholera-free, matched controls the discriminating factor was the use of a particular public well. The cases used it, while the controls did not. Subsequently, this public well was established as the source of the cholera. Similarly, the child clinical researcher, in hopes of ruling in or out the classroom context in a clinic-referred child's antisocial behavior, may draw matched controls from the target child's classmates as well as matched controls from other grade-equivalent classrooms in the child's school or neighborhood.

FUTURE DIRECTIONS AND RECOMMENDATIONS FOR TRAINING CHILD CLINICAL AND PEDIATRIC PSYCHOLOGISTS IN THE PUBLIC HEALTH AND EPIDEMIOLOGICAL PERSPECTIVES

In conclusion, we feel that future research in the child clinical and pediatric psychology arenas can benefit from the integration of life course developmental theory, community epidemiology, public health, and experimental field trials. The advantages of theory-driven research is certainly in keeping with the tradition of developmental and clinical psychology, along with the use of intervention trials to test clinical theory. The public health and community epidemiological perspectives are probably less familiar. It has been our experience that psychologists often think of epidemiological studies as simply big samples and/or community sam-

ples. We hope we have cleared up that misconception for the reader. Community epidemiology involves the careful definition of the population of interest and the context within which that population exists. The population may be as small as the attendees of an annual family reunion, which was identified as the likely context for an outbreak of food poisoning, or the entering cohort of first graders in a neighborhood elementary school. As such, unlike national probability samples, a community epidemiological approach affords the researcher the opportunity to study an individual within a community context. We also discussed two methods for integrating studies of microsocial processes into population designs: multistage sampling, which represents an efficient and economical approach to maintaining the representativeness of the individuals studied at the microsocial level within a population design, and case-control designs, which are more appropriate when the phenomena of interest is quite rare. Finally, we described the advantages of the public health perspective, which includes increased access to the populations of interest, more rapid diffusion of innovative interventions, and greater integration of those interventions into the social institutions responsible for the public's mental health. Although the integration of the community epidemiological and public health perspectives may require considerably more effort on the part of the child clinical or pediatric psychology researcher, particularly with respect to defining the population of interest and building the necessary collaborative relationships with public health institutions, we believe the product will be worth the effort.

Undoubtedly, in the absence of formal training we are unlikely to see the application of public health and epidemiological principles in child clinical or pediatric psychology research. One solution to this problem would be to require child clinical and pediatric psychology students interested in a research career to obtain a masters in public health (MPH). Most major universities with clinical-pediatric psychology training programs offer such a degree through either their schools of public health or medicine. The typical students are physicians and nurses; however, Johns Hopkins has seen of late an ever-increasing number of mental health professionals seeking an MPH. The MPH degree typically requires a 1-year commitment, with required course work and field study in epidemiological methods, biostatistics, and public health. Alternatively, Johns Hopkins offers its undergraduates as well as graduate students an opportunity to minor in public health and epidemiology. Thus, clinical-pediatric psychology students could be afforded the opportunity of taking the public health and epidemiological methods required for the MPH, but dispense with the electives or field placement requirements. Long-distance learning may be a viable alternative for those clinical-psychology training programs without a companion school of public health or medicine. Of note, each summer Johns Hopkins offers a 3-week course in the basic elements of public health and epidemiology. A number of our research psychologist colleagues have taken this 3-week course in order to make greater use of public health and epidemiological principles in their research. Minicourses such as those offered at Johns Hopkins and other major schools of public may be the best vehicle for training postdoctoral clinical and pediatric psychology researchers in the public health and epidemiological perspectives. Johns Hopkins also has postdoctoral programs in psychiatric epidemiology and prevention and mental health services

research. Fellows in each of these training programs receive training in public health epidemiology, and biostatistics.

REFERENCES

- 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: US Government Printing Office.
- Berkson, J. (1946). Limitations of the application of the fourfold contingency table. *Biometrics Bulletin*, 2, 47–53.
- Bird, H., Gould, M., Yager, T., Staghezza, B., & Canino, G. (1989). Risk factors for maladjustment in Puerto Rican children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 28, 847–850.
- Bronfenbrenner, V. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Brook, J. S., Nomura, C., & Cohen, P. (1989). A network of influences on adolescent drug involvement: Neighborhood, school, peer, and family. *Genetic, Social, and General Psychology Monographs*, 115, 125–145.
- Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues, and research strategies. *Journal of Child Psychology and Psychiatry*, 32, 1063–1080.
- Cicchetti, D., & Schneider-Rosen, K. (1984). Toward a transactional model of childhood depression. In D. Cicchetti & K. Schneider-Rosen (Eds.), *Childhood depression: A developmental perspective* (pp. 5–28). San Francisco: Jossey-Bass.
- Coie, J., Watt, N. F., West, S. G., Hawkins, D., Asarnow, J. R., Markman, H. J., Ramey, S. L., Shure, M. B., & Long, B. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist*, 48, 1013–1022.
- Crum, R. L., Lillie-Blanton, M., & Anthony, J. C. (1996). Neighborhood environment and opportunity to use cocaine and other drugs in late childhood and early adolescence. *Drug and Alcohol Dependence*, 43, 155–161.
- Goldsmith, H., Lee, A. S., & Rosen, B. M. (1984). *Small area social indicators*. Publication No. (ADM) 82-1189. Washington, DC: US Government Printing Office.
- Gordis, L. (1996). *Epidemiology*. Philadelphia, PA: W. B. Saunders.
- Greenley, J. R., & Mechanic, D. (1976). Social selection in seeking help for psychological problems. *Health and Social Behavior*, 17, 249–262.
- Greenley, J. R., Mechanic, D., & Cleary, P. (1987). Seeking help for psychologic problems: A replication and extension. *Medical Care*, 25, 1113–1128.
- Ialongo, N., Reider, E., & Kellam, S. (submitted). A further look at the prognostic power of young children's reports of depressed mood and feelings.
- Ialongo, N., Werthamer, L., Brown, H. B., Kellam, S., & Wai, S. B. (in press). Proximal impact of two first grade preventive interventions on the early risk behaviors for later substance abuse, depression, and antisocial behavior. *American Journal of Community Psychology*.
- Jason, L. A. (1982). Community-based approaches in preventing adolescent problems. *School Psychology Review*, 11(4), 417–424.
- Johnston, L. D., O'Malley, P., Bachman, J. G. (1995). *National survey results on drug use from the monitoring the future study, 1975–1994: Vol. 1, Secondary School Students* (DHHS, PHS, NIH, Publication No. 95-4026). Washington, DC: US Government Printing Office.
- Kellam, S. G. (1990). Developmental epidemiologic framework for family research on depression and aggression. In G. R. Patterson (Ed.), *Depression and aggression in family interaction* (pp. 11–48). Hillsdale, NJ: Lawrence Erlbaum.
- 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: University of Chicago Press.
- Kellam, S. G., Branch, J. D., Brown, C. H., & Russell, G. (1981). Why teenagers come for treatment: A ten-

- year prospective epidemiological study in Woodlawn. *Journal of the American Academy of Child Psychiatry*, 20, 477–495.
- Kellam, S. G., & Ensminger, M. E. (1980). Theory and method in child psychiatric epidemiology. In F. Earls (Ed.), *Studying children epidemiologically*. International Monograph Series in Psychosocial Epidemiology, Vol. 1 (pp. 145–180). New York: Neale Watson Academic Publishers.
- Kellam, S. G., Ensminger, M. E., & Turner, R. J. (1977). Family structure and the mental health of children: Concurrent and longitudinal community-wide studies. *Archives of General Psychiatry*, 34, 1012–1022.
- Kellam, S. G., & Hunter, R. C. (1990). Prevention begins in first grade. *Principal*, 70, 17–19.
- Kellam, S. G., Ling, X., Merisca, R., Brown, H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology*, 10, 165–185.
- Kellam, S. G., Mayer, L. S., Rebok, G. W., & Hawkins, W. E. (in press). The effects of improving achievement on aggressive behavior and of improving aggressive behavior on achievement through two prevention interventions: An investigation of etiological roles. In B. Dohrenwend (Ed.), *Adversity, stress and psychopathology* (pp. 486–505). New York: Oxford University 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: Guilford Press.
- Kellam, S. G., Rebok, G. W., Mayer, L. S., Ialongo, N., & Kalodner, C. R. (1994a). Depressive symptoms over first grade and their response to a developmental epidemiologically based preventive trial aimed at improving achievement. *Development and Psychopathology*, 6, 463–481.
- Kellam, S. G., Rebok, G. W., Ialongo, N., & Mayer, L. S. (1994b). 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 and Allied Disciplines*, 35, 259–281.
- Kellam, S. G., & Werthamer-Larsson, L. (1986). Developmental epidemiology: A basis for prevention. In M. Kessler & S. E. Goldston (Eds.), *A decade of progress in primary prevention* (pp. 154–180). Hanover, NH: University Press of New England.
- Leaf, P., Alegria, M., & Cohen, P. (1996). Mental health service use in the community: Results from the four community MECA study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 889–897.
- Lilienfeld, D. E., & Stolley, P. D. (1994). *Foundations of epidemiology*, 3rd ed. New York: Oxford University.
- Minkler, M. (1990). Improving health through community organization. In K. Glanz, F. M. Lewis, & B. Rimer (Eds.), *Health behavior and health education: Theory research and practice* (pp. 257–287). San Francisco: Jossey-Bass.
- Mirsky, A. F., Anthony, B. J., Duncan, C. C., Brouwers, P., Ahearn, M. B., & Kellam, S. G. (1991). Analyses of the elements of attention: A neuropsychological approach. *Neuropsychology Review*, 2, 109–145.
- Morris, J. N. (1975). *Uses of epidemiology* (3rd ed.). New York: Churchill Livingstone.
- Mrazek, P. G., & Haggerty, R. J. (Eds.). (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.
- National household survey on drug abuse: Population estimates 1994* (1995). U.S. DHHS, PHS, DHHS Pub. No. (SMA)95-3063. National health interview survey: Form HIS-1A (1998). U.S. DHHS, PHS.
- National Institute of Mental Health. (1991). *Implementation of the national plan for research on child and adolescent mental disorders* (Program Announcement PA-91-46). Washington, DC: US Government Printing Office.
- Offord, D. R., Boyle, M. H., & Racine, Y. A. (1989). Ontario Child Health Study: Correlates of Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 28, 856–860.
- Offord, D., Boyle, M., & Szatmari, P., (1987). Ontario Child Health Study II: Six-month prevalence of disorder and rates of service utilization. *Archives of General Psychiatry*, 44, 832–836.
- Pentz, M. A. (1993). Directions in the future research in drug abuse prevention. *Preventive Medicine*, 23, 646–652.
- Perry, C., Williams, C., Veblen-Mortenson, S., Toomey, T., Komro, K., Anstine, P., McGovern, P., Finnegan, J.R., Forster, J., Wagenaar, A., & Wolfson, M. (1996). Project Northland: Outcomes of a community-wide alcohol use prevention program during early adolescence. *American Journal of Public Health*, 86, 956–965.

- Reid, J. (1993). Prevention of conduct disorder before and after school entry: Relating interventions to developmental findings. *Development and Psychopathology*, 5, 243–262.
- Sandler, I. N., West, S. G., Baca, L., Pillow, D. R., GerstOn, J. C., Rogosch, F., Virdin, L., Beals, J., Reynolds, K. D., Kallgren, C., Tien, J., Krieg, G., Cole, E., & Ramirez, R. (1992). Linking empirically based theory and evaluation: The family bereavement program. *American Journal of Community Psychology*, 20, 491–521.
- Schlesselman, J. (1981). *Case-control studies: Design, conduct, analysis*. New York: Oxford University Press.
- Snow, J. (1849). On the pathology and mode of communication of cholera. *London Medical Gazette*, 44, 745–752, 923–929.
- Tweed, D. L., Goldsmith, H. F., Jackson, D. J., Stiles, D., Rae, D. S., & Kramer, M. (1990). Racial congruity as a contextual correlate of mental disorder. *American Journal of Orthopsychiatry*, 60, 392–402.
- Wagenaar, A., Murray, D., Wolfson, M., Forster, J., & Finnegan, J. R. (1994). Communities mobilizing for change on alcohol: Design of a randomized community trial. *Journal of Community Psychology, CSAP Special Issue*, 79–101.
- Zaner, G., Pawelkiewicz, W., DeFrancesco, J., & Adnopo, J. (1992). Children's mental health service needs and utilization patterns in an urban community: An epidemiological assessment. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 951–960.

2

Testing Theoretical Models and Frameworks in Child Health Research

THOMAS A. WILLS and SEAN D. CLEARY

INTRODUCTION

In this chapter we discuss testing theoretical models of processes in child health research. By this we mean the kind of research that is based on deriving predictions from a theoretical portrayal of the process that has engaged the investigator's interest and designing a study to provide a test of this model. Some may think that theoretical models are always complicated and abstruse, but this is not the case; in fact, some of the best models may be quite simple ones. A child psychologist may pose a question such as, "Why are some children more at risk for a certain condition?" or "How do families adapt successfully to their child's chronic disease?" or "What makes a particular treatment technique effective?" The psychologist's thinking about the process underlying the outcome provides the basis for a model of how things occur: How do environmental and familial factors combine to create risk; what coping processes lead to adaptation; what mediating variables are responsible for the effectiveness of a therapeutic program. Such statements are the beginning of a testable model.

A theoretical model may use suggestions from clinical experience with children and families, may combine known and unknown elements based on findings from previous research, or may derive predictions from an existing theoretical

THOMAS A. WILLS and SEAN D. CLEARY • Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, Bronx, New York 10461.

statement.* The model then provides the researcher with additional leverage for studying his or her question of interest. If the results are consistent with prediction, this strengthens confidence in the meaningfulness of the model and suggests a larger picture of the processes producing risk or adaptation, which may go beyond a particular disease condition or clinical population. If the findings are in some respects inconsistent with prediction, the results usually suggest ways in which the theory may be modified so as to be in better accord with clinical reality. In either event, the researcher has learned something and has a springboard for design of further research that will provide a better understanding of children's health. Thus, the advantage of theory-driven research is that it can sharpen the investigator's thinking about a question and provide guidance to others who are interested in studying the question.

The purpose of this chapter is to consider the tools used in theory-driven research. While some of these may already be familiar to the reader, we will "walk through" the steps along the way, starting with derivation of hypotheses and going through the design of a study and testing of a model. The chapter will consider several statistical approaches, including structural modeling procedures, discussed in a nontechnical manner. We will assume that the reader has an understanding of basic statistical terminology and methods and is interested in learning more about what structural modeling procedures can do. We do not intend to address all the technical details of the analytic procedures; these are covered in books written at introductory and advanced levels (see, e.g., Hayduk, 1987; Hoyle, 1995; Schumacker & Lomax, 1996). We also cannot cover all possible theories; instead, our goal is to present a range of models that are represented in current research on child clinical psychology.

In the following sections, we first review different types of theoretical approaches and discuss the kinds of predictions they can generate. Then we consider several types of statistical models that may be appropriate for testing different types of questions. The next sections demonstrate how to construct and test predictions in multiple regression and structural modeling. The process of specifying and testing a model will be illustrated with examples from the literature on pediatric psychology.

WORKING FROM A THEORY

The reader is probably familiar with the general definition of theory as a set of postulates that are internally consistent and attempt to provide an explanation for an observable phenomenon. The postulates may be used to explain or integrate previously observed phenomena; to suggest research designs that can be used to test the validity of the theory in accounting for such phenomena; and/or to predict new phenomena that have not yet been observed. What is less obvious is that theories vary tremendously in level of specificity versus generality. In the area of substance use, for example, there are psychopharmacological models that try to account for the paradoxical effects of nicotine (Gilbert, 1979; Perkins, 1995), theo-

*A theory may start with simple clinical observation and develop in richness and complexity as research proceeds. An example of theory derived from clinical observation includes Type A theory, based on the observation that chairs wore out faster in waiting rooms from cardiac patients, who were later described as restless and irritable (Friedman & Rosenman, 1974).

retical models of how tobacco and alcohol may be involved in coping with stress (Khantzian, 1990; Wills & Shiffman, 1985); and theories that address the entire range of environmental, individual, and social factors that are related to adolescent substance use (Petratatis, Flay, & Miller, 1995; Wills, Pierce, & Evans, 1996). None can be said to be better than the others, for their goals are different. For example, the most specific (pharmacological) model has been useful for research on relapse episodes; the coping models have been utilized in research on problematic substance use; and the most general models have been applied in epidemiological research on the distribution of risk factors and substance use within the general population. Accordingly, a researcher's gravitation toward a particular type of theoretical structure would be shaped by his or her research interests and the overall direction of the research program.

Types of Theories

In the domain of child health research, several types of theories may be relevant. Briefly we can outline the essential propositions of five types of theories. We refer to related groups of theories because in most areas there are several variant models that contain common sets of propositions.

Social Influence Models

Because the social environment is an important influence on behavior, some models aim to describe the processes through which peers, parents, or other social agents have an impact on children's health-related attitudes and behavior. These models usually acknowledge Bandura's social learning theory as their intellectual progenitor (Bandura, 1969, 1977). The aim of social influence models is to explain how social agents influence behavior. Accordingly, they focus attention on particular variables, such as the attitudes and behavior of peers and parents, and describe processes through which social factors may operate, such as identification, modeling, or overt social pressure. In addition, these models may call attention to particular situations in which social influences are prominent, such as peer group initiation of smoking (Biglan, Weissman, & Severson, 1985; Friedman, Lichtenstein, & Biglan, 1985). Social influence models provide specific hypotheses about which social agents will be most important as potential influences on children's behavior and they suggest assessment approaches using questionnaire measures or observational approaches to understand how social influences operate. Such models include cognitive-behavioral concepts and have drawn attention to how children's perceptions of social norms may influence behavior (Graham, Marks, & Hansen, 1991) and may be relevant for prevention programs (Hansen & Graham, 1991). Inclusion of developmental concepts also has suggested how peers and parents may have differential influence for children at different ages (e.g., Flay et al., 1994).

Family Systems Theories

Family systems models have examined how interrelationships among family members result in certain outcomes. Using the family as the basic unit of analysis, these models suggest ways of conceptualizing the set of relationships among

parents and children (e.g., Doherty & Campbell, 1988; Ramsey, 1989). Some models are more developmental in nature, emphasizing the importance of the family milieu (Hauser, 1991; Hauser et al., 1990). Others have used psychodynamic concepts to suggest propositions about how a life change for one family member will disrupt the current balance of relationships and result in a new configuration of family interaction, with different consequences than before (Minuchin, 1974; Turk & Kerns, 1985). Family systems models suggest approaches to assessing complex aspects of family relationships through analogue procedures as well as clinical interviews. Some applications have considered how family relationships may support health behavior change in some situations but may operate to undermine behavior change in other situations (Baranowski & Nader, 1985).

Stress and Coping Theories

Models derived from the general theory of stress and coping processes (Lazarus & Folkman, 1984) have considered how individual coping responses affect adaptation to an accumulation of stressful life events. These models suggest dimensions of stressors that are most salient and types of coping responses that may be adaptive or maladaptive for certain types of stressors; in addition, they suggest how coping may change over time as various aspects of a stressor are resolved (Folkman & Lazarus, 1985). Some models have focused on how stress and coping are implicated in the development of children's behavior problems (Compas, 1987; Compas, Malcarne, & Banez, 1992; Wills, 1986; Wills & Filer, 1996). Because having a child with a chronic medical condition or behavioral problem is an enduring stressor, coping models also have been developed for understanding families' adaptation in these situations (e.g., Carver, Scheier, & Pozo, 1992; Cole & Reiss, 1993; Thompson et al., 1994; Wallander, Varni, Babani, Banis, & Wilcox, 1989). Stress and coping models generally draw attention to types of life events that can disrupt children's or families' adaptation and how appropriate coping responses can help to restore adequate functioning.

Buffering Models

These models originated from the observation that persons with good social support were less affected by negative life events. From this came the concept that social support reduces (buffers) the potential impact of adverse conditions (Cohen & Wills, 1985). The buffering effects of social support have been demonstrated for a range of dependent variables, including psychological symptomatology, physical health, and substance use (Wills & Filer, 1999). Buffering models are conceptually related to theories of resiliency effects, which derived from observations that some children growing up in adverse circumstances (e.g., poverty, parental substance abuse) showed relatively good functioning at later ages (Garnezy & Masten, 1991; Rutter, 1990; Werner, 1986). Resiliency models have focused attention on characteristics of children, such as temperament, social skills, and academic competence, that may serve as resilience factors (Masten, Morison, Pellegrini, & Tellegen, 1990; Wills, Blechman, & McNamara, 1996). Buffering and resiliency models draw attention to social resources or individual characteristics that may help to counter the impact of adverse circumstances and to the possible mechanism of how these effects occur (Dubow & Tisak, 1989; Rosenbaum, 1990; Wills & Cleary, 1996).

Diathesis-Stress Models

These types of models inquire how particular vulnerability factors may predispose some children to a disorder and how the latent vulnerability (i.e., diathesis) may be evoked by certain types of circumstances. Particularly relevant for child health research are models of how family history of depression may make children more vulnerable to emotional problems (Monroe & Simons, 1991), and how a family history of alcoholism may make children more vulnerable to substance abuse (Sher, 1991; Windle & Searles, 1990; Zucker, 1994). Diathesis-stress models focus attention on family characteristics that increase vulnerability for a given disorder and may suggest something about the mechanism of how this occurs. However, these are not mutually exclusive of models that suggest protective factors that may help to offset vulnerability, and recent research includes examples of combined models that examine how children's functioning is shaped through the balance of vulnerability and protective factors (e.g., Newcomb & Felix-Ortiz, 1992; Wills, Vaccaro, & McNamara, 1992).

In summary, we have outlined the nature of several theoretical models that may have some relevance for child health research. Reading of theoretical papers may be helpful for a clinical researcher in the early stages of research development. A theoretical model may be useful for guiding the researcher's thinking about the nature of a process, suggesting choices of variables or generating predictions. It is of course impossible to advocate a particular theory, because as noted earlier the goals of theoretical models are so different. The only clear recommendation we can make is "read what makes sense to you" and then go from there. Thoughts generated from prior theoretical papers may help a researcher to see certain directions as more promising, to reject other directions as irrelevant, or to add variables and mechanisms that had not been included previously in his or her thinking. Alternatively, the researcher may conclude that existing theory is misguided and then articulate a new model based in his or her own thinking and clinical experience.

THE NATURE OF THE MODEL

A theoretical model provides a conceptual base for design of the research, suggesting what variables can be measured and how the nature of a process can be conceptualized. This line of thinking then leads to the specification of an analytic model that should represent the process in question.* A number of models are available to test a hypothesized process, the choice of the model being suggested by the nature of the researcher's question.

Types of Analytic Models

Direct-Effects Models

We use an unfamiliar term to introduce several procedures that are probably familiar to the reader: univariate regression and multiple regression. The purpose

*We say "analytic model" rather than statistical model because a given question may be tested through several statistical methods; for example, simple mediation hypotheses do not necessarily require structural modeling procedures but may be tested in multiple regression or even with simple partial correlations, which can be computed by hand.

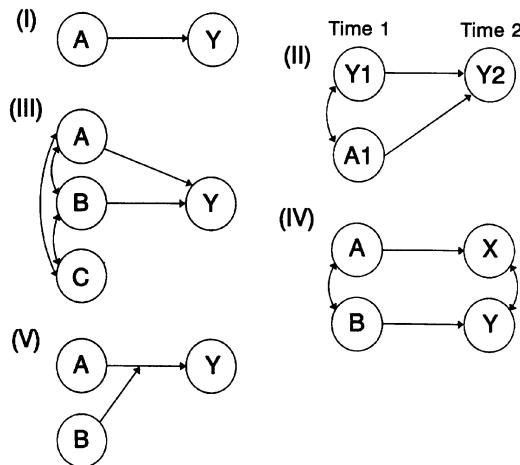


Figure 1. Direct effect models: (I) simple effect model; (II) prospective simple effect model; (III) multiple effects model; (IV) multiple effects model with two or more criterion variables; and (V) interaction model.

is to point out that these models involve an important but usually unstated assumption: the predictor is directly related to the outcome, without any intermediate process(es) being involved.

Simple Effect Model. A simple effect or bivariate model (Fig. 1-I) contains a single predictor (*A*) and a single outcome (*Y*) variable (see Appendix for a glossary of relevant terms and concepts). This is an example of a simple model to test whether *A* is related to *Y*. A significant effect in either direction, positive or negative, indicates the variables are related. The model represents a hypothesis that can be tested using simple linear regression, without consideration of the impact of other variables. The longitudinal extension (Fig. 1-II) involves assessment of variables at two or more time points. Multiple regression is used to predict the criterion variable at time 2 (*Y*2) from the predictor at time 1 (*A*1), including the level of the predictor at time 1 as a predictor, which now would technically be termed a covariate (*Y*1). The finding that *A* is a significant predictor of *Y*2 is equivalent to the statement that it predicts change over time in the criterion, above and beyond the temporal stability of the variable (*Y*1-*Y*2), which typically is substantial. Longitudinal results have the advantage of demonstrating temporal ordering in the relationship, that is, “level of *A* precedes change in level of *Y*,” and the investigator may then conclude in some kind of language that “*A* causes *Y*,” although a causal conclusion is not strictly warranted.* What we are emphasizing here is that

*Note that in most cases, multiple regression or structural modeling procedures do not strictly allow inferences about causality because they are based on correlational data. In this discussion, however, it is conventional to use language suggesting that one variable causes another, because this is how models are specified.

whether cross-sectional or longitudinal, this model assumes there is a direct relationship between A and Y .

Multiple Effects Model. It is possible that several variables are anticipated to make independent contributions to an outcome; for example, current family interaction and a child's previous treatment history might each predict a clinical outcome and not be redundant with each other for predicting the outcome. In this case, a multiple effects model is specified (Fig. 1-III). There are two differences from the previous model. The first is that model III posits that the outcome in question is determined by several different predictive factors. This model is plausible in real-life situations, where a child's functioning is frequently determined by many factors. The second issue is that in model III the analysis accounts for any correlation between the predictors (A and B), again a common situation in clinical research as predictor variables are often correlated. The analytic power in model III is the demonstration that several predictor variables make significant unique contributions to an outcome even when the intercorrelation of the predictors is statistically removed. Possible confounders (i.e., third variables that are correlated with a predictor and the outcome) may be included in multivariate analysis to determine whether they are partially or completely responsible for the observed correlation between A and Y , as indicated in the figure by C . Demonstration of a predictive relationship with control for possible confounders (e.g., family's socioeconomic status, child's gender) increases confidence in the results. The longitudinal extension of the model, involving prediction of Y at time 2 from the predictors at time 1 and the covariate at time 1, is also a useful procedure but is not diagrammed. Another type of model involves two or more criterion variables, as in Fig. 1-IV. Here, the relation of two predictor variables (A and B) to two different criterion variables (X and Y) is tested; the criterion variables are specified as correlated. This model can be tested with multivariate analysis of variance (MANOVA) or multivariate multiple regression (Cohen & Cohen, 1983, chap. 12; Stevens, 1996), but is most straightforward in structural modeling (Jöreskog & Sörbom, 1988).

Interaction Model. When a buffering or resiliency process seems plausible, the appropriate analytic model is one that tests for a moderation effect; this is diagrammed in Fig. 1-V. One predictor (A) is hypothesized to have a different effect on the criterion (Y) at different levels of the other (B); for example, negative life events could have less effect on depression for persons with a high level of social support. The buffering hypothesis can be tested as an interaction effect in analysis of variance or in multiple regression by entering the two predictors together with their cross-product (Cohen & Cohen, 1983). What is not generally appreciated is that tests of interaction effects have different power requirements than tests of main effects (Aiken & West, 1991; McClelland & Judd, 1993). It is not uncommon to see studies testing buffer effects with designs that had little likelihood of detecting such interactions, and researchers who are serious about testing interaction models have to consider increased sample size or stratified designs in order to have reasonable power (McClelland & Judd, 1993). Diathesis-stress theory also implies an interaction model, as in Fig. 1-V. However, in this case the stress factor is predicted to have more effect when the diathesis is present, that is, persons are more vulnerable to life stress. The statistical procedures are the same, but the

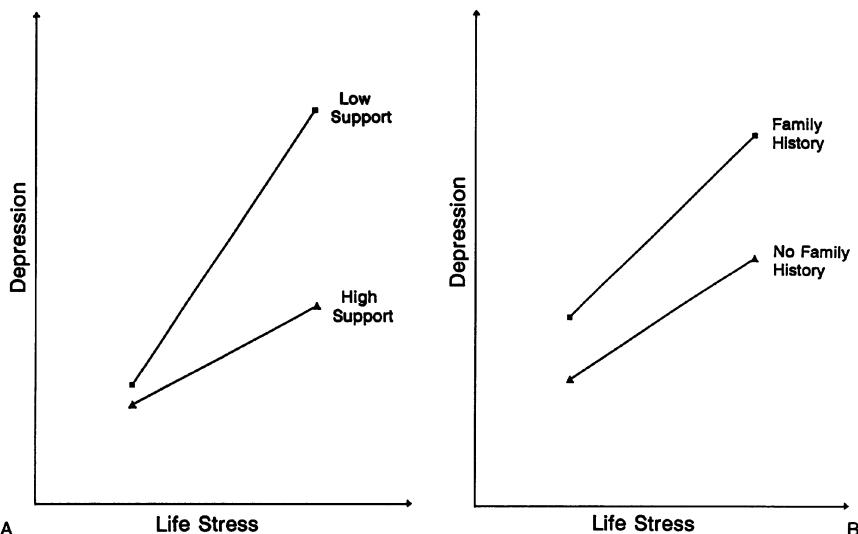


Figure 2. Interaction effects: (a) stress-buffering effect; and (b) diathesis-stress effect.

nature of the interaction is different. Examples of interaction effects are presented in Fig. 2. Figure 2A illustrates a stress-buffering effect, showing that the relationship between life stress and depression is reduced for persons with high support; Fig. 2B illustrates a diathesis-stress effect, showing that life stress is strongly related to depressive symptoms among persons with a family history of depression, but has less effect for persons without such a history.

Mediation Models

Mediation models involve an important difference in assumptions about the nature of the process under observation. Some interesting questions in pediatric psychology involve the concept that the effect of one variable on an outcome is transmitted (mediated) through another variable. For example, the protective effect of family support on children's depressive symptomatology could be mediated through increasing the child's adaptive coping (Thoits, 1986; Wills & Filer, 1996). While such models can be tested in multiple regression, they involve a basic difference in the conceptualization of the process.

Simple Effect Mediation Model. A simple effect hypothesis might state, for example, that the relation of family communication to children's depressive symptomatology is mediated through children's beliefs about the world (see Stark, Schmidt, & Joiner, 1996). If this hypothesis is true, the data will show a modest correlation between family communication and depression, but a stronger correlation between communication and children's belief structure and a stronger correla-

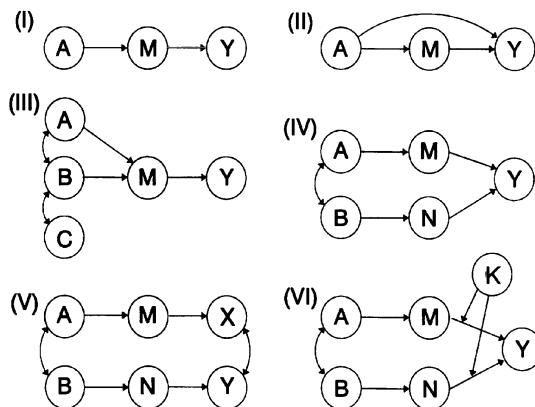


Figure 3. Mediation models: (I) complete mediation model; (II) partial mediation model; (III) multiple effects mediation model; (IV) multiple mediator model; (V) multiple mediator model with two or more criterion variables; and (VI) multiple mediator model with moderation.

tion between children's beliefs and depressive symptomatology. Procedures for testing correlations for evidence of mediation are described in several places (Baron & Kenny, 1986; Holmbeck, 1997). The analytic model is represented in Fig. 3-I, which shows how the effect of predictor *A* on criterion *Y* can be transmitted through a mediator variable *M*. The mediation hypothesis can be tested in multiple regression first by analysis with *M* predicted from *A* alone, and second by analysis with *Y* predicted from *A* and *M* entered together. There are two possible results from such an analysis. In a complete mediation process, the effect of *A* is entirely transmitted through *M*; this is diagrammed in Fig. 3-I. In a partial mediation process, some of the effect of *A* is transmitted through *M*, but there is also evidence of a direct effect, a significant effect of *A* on *Y* even when entered with *M*; this is diagrammed in Fig. 3-II.* An extension that is easily testable in multiple regression is diagrammed in Fig. 3-III. Here, the investigator suspects that effects of several predictor variables may be mediated through an intermediary (*M*) and specifies a model providing a comparative test, which not only controls for the correlations among the predictors but also tests whether each of the predictors (*A*, *B*, and *C*) is mediated through *M*. Simple effect models can be used in predictive research, for example testing how the relation between parental substance use and adolescents' substance use is mediated (Wills, Schreibman, Benson, & Vaccaro, 1994). They also are useful for intervention research, as they can be used to test hypotheses about how the effect of a prevention program or treatment program is mediated (Hoyle & Smith, 1994; MacKinnon et al., 1991; MacKinnon, 1994).

*There is a third possibility, in which there is only a direct effect of *A* on *C*; in this case, one would say there is no evidence of mediation. In a sense this is a test of an alternative model, i.e., for a mediated versus nonmediated effect, so one could use this analytic model for comparative tests. Baron and Kenny (1986) outline a set of procedures to test zero-order correlations for evidence of mediation, so the researcher will have preliminary evidence on whether mediation exists.

Multiple Mediator Models. What if the effects of predictors are mediated through different intermediary variables? One such model is outlined in Fig. 3-IV. In this situation the researcher hypothesizes that the effect of predictor *A* is mediated through a different factor (*M*) than is true for predictor *B*, whose effect is mediated through another intermediary (*N*). Such models offer potential advantages for clinical research because they allow several mediation hypotheses to be tested in a single study, as long as the relevant predictor and mediator variables are included. (Direct effects are also possible in such models though not included in the example.) An extension of this model that includes two or more related criterion variables (*X* and *Y*) with multiple predictors (*A* and *B*) that are mediated through different factors (*M* and *N*) is displayed in Fig. 3-V. Multiple mediator models can be analyzed in multiple regression, but for complex models it is more feasible to use structural modeling procedures such as LISREL (Jöreskog & Sörbom, 1988) or EQS (Bentler, 1989). When expert consultation is available, models with high levels of complexity can be analyzed.

Mediation with Moderation. In outlining both simple and multiple mediator models, we have described procedures for understanding what would be termed main effect relationships in the mediation process; it is assumed that the mediation effects are not affected by other factors. However, it is possible that an interaction (i.e., moderation effect) exists, so that the magnitude of mediation paths differs by the level of another variable. For example, it is possible that the effects of social support are different for persons with low versus high life stress. Such a question can be addressed with what are termed multiple group (or stacked) procedures in structural modeling, where mediation models are analyzed for subgroups that differ on another variable (Jaccard & Wan, 1996). An example is diagrammed in Fig. 3-VI. This example represents mediation because the effects of *A* and *B* are mediated through other variables *M* and *N*; it also includes moderation because the magnitudes of the mediator effects (path *M* → *Y* and path *N* → *Y*) differ according to the level of a third variable *K*. Such procedures can be employed to test buffering hypotheses in a multivariate context; for example, one study tested a multiple mediator model for effects of family support to determine where buffering effects occurred (Wills & Cleary, 1996).

CONSTRUCTING A MODEL

Consider the situation where a person's clinical experience, reading, and thinking have produced a theoretical position and a set of hypotheses. From this point, the researcher is in a position to design and conduct the study and to analyze the data. The construction of an analytic model is inherently guided by the clinical researcher's thinking about the processes that combine to produce an outcome; in fact, they are the same thing. The conceptualization of the process is the same thing as the design of the study that is conducted, which is the same thing as the statistical model that is analyzed to yield the results. To paraphrase Gertrude Stein, the theory is the study is the model.

Here we describe the conceptual tools used in constructing a model. As will become evident, these cover the design of the study as well as the initial construc-

tion (technically termed *specification*) of a statistical model. We introduce terminology for concepts that may be unfamiliar to some readers; this is done to illustrate how these concepts are used in constructing a model and also may be useful for understanding technical terms used in the literature. Note that in the following section we shift from talking about “variables” to talking about “constructs,” because in structural modeling procedures a theoretical construct can be indexed by several different indicators and it is the construct that is actually analyzed.

Distal and Proximal Factors

An important difference from regression research, with a predictor and a criterion, is the concept that there are separate theoretical levels of predictors. The distinction between distal and proximal factors was developed to address the case where an outcome has multiple determinants that are themselves related (Jessor & Jessor, 1977). It is assumed that the outcome, say, antisocial behavior, occurs through several variables that are linked in a causal chain. A distal factor is one that is posited to be far back in the causal chain, that is, relatively remote from the actual behavior; a proximal factor (from the Latin *proximus*, nearest to) is one that is posited to be the closest causal factor to the occurrence of the behavior. A proximal factor is the final step to the behavior. It is not identical to the behavior, but a high level of a proximal factor greatly increases risk for the behavior (Moffitt, 1993).

For example, problem behavior theory (Jessor & Jessor, 1977) posits that adolescent problem behavior is rooted in distal factors such as poor parent-child relationships; that children who are rejected by parents disengage from conventional values and form attitudes more tolerant of deviant behavior; and that younger adolescents with deviance-prone attitudes will begin to associate with peers who hold similar attitudes. Affiliation with deviant peers is posited to be the proximal factor for deviant behavior, because a deviant peer group may then express attitudinal support for deviant behavior, provide the physical means for the behavior (e.g., getting beer from an older brother), or provide the organizational means to accomplish it (e.g., extorting money from other students).

In this system, the parent-child relationship is assumed to be a relatively distal factor, deviance-prone attitudes are posited to have an intermediate status, and affiliation with deviant peers is the proximal factor for the occurrence of the behavior (cf. Patterson, DeBarryshe, & Ramsey, 1989). Note that this theory does not posit an invariant linkage of constructs, because a poor parent-child relationship may be offset by other factors, such as good relationships with other persons (e.g., siblings, relatives, or teachers). However, the prediction is that other things equal, children having a poor relationship with parents will be more likely to develop tolerance of deviance; and other things being equal, child holding deviance-prone attitudes will be more likely to affiliate with deviant peers. The theory holds that the proximal factors tend to follow from the distal factors. This type of theory is also represented by Moffit's model on neuropsychological antecedents of conduct disorder, which has proposed how early, subtle neuropsychological deficits such as verbal fluency may be linked to later, more proximal risk factors for antisocial behavior (Moffitt, 1993; Moffitt & Lynam, 1994) and by models that suggest how parents' low socioeconomic status is linked to adverse outcomes for children (Elder, Nguyen, & Caspi, 1985; Wills, McNamara, & Vaccaro, 1995).

Specifying Exogenous and Endogenous Constructs

In model-testing procedures, an exogenous construct is defined as something that is not caused by any other construct in the model; an endogenous construct is defined as something that is caused by a prior construct in the model. For example, in Fig. 1-I it could be said that A is exogenous and Y is endogenous, and this seems like a trivial restatement. However, in Fig. 3-I the terminology is more useful, because there is one exogenous construct and two endogenous constructs; the model is specified to test the proposition that A (the exogenous construct) causes M (the first endogenous construct), and that M causes Y (the second endogenous construct). Extension of this terminology in Fig. 3-III is that there are three exogenous constructs (A , B , and C) and two endogenous constructs (M and Y); and in Fig. 3-IV, that there are two exogenous constructs (A and B) and three endogenous constructs (M , N , and Y). In structural equation models, exogenous variables appear on the left side of the model with arrows from the variable to one or more endogenous constructs, indicating the theoretical proposition that in some sense they cause the endogenous construct(s). Note that in Figs. 3-III and 3-IV, there are curved arrows among the exogenous constructs, to indicate that the correlations among these constructs are included as parameters in the analysis; exogenous constructs are specified as correlated because (following from the definition) they are not caused by any prior construct in the model.

In direct effect regression procedures (Fig. 1) there is no theoretical discretion about exogenous and endogenous constructs; there is a set of predictor variables (exogenous) and a criterion variable (endogenous), and that is the end of the story. Theory-testing procedures in structural modeling, however, involve more thinking about the specification of the model. A construct may be specified as exogenous because it is logically and temporally prior to another construct. For example, in a study considering the relation between gender and behavior problems, gender would be exogenous because logically it exists prior in time to the onset of behavior problems. Relevant control constructs would typically be specified as exogenous constructs because they are assumed to be correlated with parts of a causal process but are not themselves causally involved in the process. Demographic variables (e.g., ethnicity) are often specified in this manner, to show that the elements in a postulated model occur independent of demographic characteristics, though demographics may be related to some of the endogenous variables in the model.

The ordering of endogenous constructs is based on logical considerations, theoretical propositions about their places in a process, and available empirical evidence about the temporal nature of relationships. For example, in a model investigating the effect of demographic variables on IQ and delinquency, IQ logically would be specified before a measure of delinquent behavior because theoretically it existed in some form before the child began to manifest delinquent behavior. By extension, another study might investigate oppositional defiant behavior as a precursor of conduct disorder; in this case the oppositional defiant construct would be specified before the conduct disorder construct because it is manifested at earlier ages than conduct disorder is manifested. The general type of the investigator's thinking in constructing a model with several endogenous variables is that theoretically, " A comes before B , and B comes before C ."

Another way of stating the principle is that the model is constructed with a progression from distal to proximal factors. With reference to the models in Fig. 3, constructs that are far back in a causal chain are exogenous, hence, go toward the left-hand side of a model, whereas constructs that are more proximal to the outcome go toward the right-hand side of the model. For example, several investigators have considered how socioeconomic status is related to adolescent problem behavior (Conger et al., 1990; Takeuchi, Williams, & Adair, 1991). In these studies, low socioeconomic status of the parents is a distal construct because it exists from a young age of the child; lower emotional support from parents as a consequence of their economic stress is posited as an intermediate endogenous construct; and increased affiliation by the adolescent with deviant peers is posited as the final endogenous construct, the proximal factor for the adolescents' behavior problems.

Construction of a model with exogenous and endogenous variables uses a combination of the following: theory and logic, prior research findings, and hypothesized relationships. Prior research can be quite useful for demonstration that relationships among several constructs exist and the patterning of relationships is coherent with the investigator's model. Prior longitudinal studies, when available, can provide evidence about temporal ordering. We note that for theoretical models postulating relationships among several mediators, it is desirable to obtain longitudinal data. However, the ideal study for a mediational model with n stages would contain $n + 1$ waves; so even a relatively simple model such as Fig. 3-III ideally would have four waves of data. Since few investigators have conducted four- or five-wave studies, it is more feasible to use a combination of theory and empirical evidence in specifying a model. The important thing to note is that once specified, a model has a highly desirable philosophical property: *It is easily falsified*. Any of the models in Fig. 3 is easily falsified if data show that predicted paths are not obtained or the model does not provide a good representation of the empirical relationships among the variables. If the model shows results consistent with prediction, the researcher's theory gains greatly in confidence, because not only do relationships exist among particular variables, but the total pattern of interrelationships is also consistent with prediction.

Specifying the Measurement Model

Another important difference from regression research involves the measurement of constructs. In multiple regression, a given variable is measured and entered in the regression as the predictor; it is recognized that the measured variable is not a perfect predictor, but the analysis can do nothing about this. In contrast, structural modeling allows situations where constructs can be measured by multiple indicators. One advantage of this procedure is that when the indicators of a theoretical construct are well selected, the reliability of measurement is increased. In addition, structural modeling includes a set of procedures termed *confirmatory analysis* that allow a statistical test of whether the theory underlying the selection of indicators is correct. Hence, structural modeling provides a "double whammy," because the investigator can test propositions about the measurement of the constructs as well as testing predictions about the relations among the constructs.

The term *measurement model* describes the procedures used to measure the constructs in the research. A distinction is made between manifest variables and

latent constructs. The term *manifest variable* describes the familiar situation where one variable is used to measure a construct, for example, a 10-item score for closeness to parents is used to measure the construct "closeness to parents." In this case, the observed variable and the construct are the same thing. In contrast, structural modeling allows a procedure where several different but theoretically related indicators are used to measure an unobserved latent construct. A latent construct must be measured by at least two manifest variables, but three or more are generally recommended. For example, a researcher could index a latent construct "closeness to parents" by using measured variables of (1) perceived emotional support from parents, (2) time spent with parents versus peers, and (3) a rating of positivity of relationship with parents as indicators. Note that the latent construct—closeness to parents—is not directly observed but is inferred from the measured values of the three indicators; typically the latent construct would then carry more information than was provided by any single indicator. One advantage of this procedure is that when indicators are appropriate, the error of measurement for the construct is greatly reduced, so other things being equal the investigator is more likely to detect a relationship that truly exists. We note that manifest variables and latent constructs may be included in the same structural modeling analysis. It is possible to include manifest variables when there is only one conceivable indicator of a construct (e.g., child's age) or when research priorities mandate that single indicators be obtained for some constructs because of limited time for data collection. There might be some loss in reliability, but there is no mathematical reason that excludes a model with both manifest variables and latent constructs.

In tests of hypotheses with structural equation modeling, the measurement model is the basis on which the structural model (i.e., relationships between constructs) is specified. The procedure termed *confirmatory analysis* provides an empirical evaluation of the measurement model and typically is the first step in structural modeling. The matrix of actual covariances among all indicators is input to the analysis, the researcher provides a specification of the constructs that the indicators are supposed to measure, and the analysis provides an index (termed a fit index) indicating the extent to which the hypothetical model successfully reproduces the actual covariances. The confirmatory procedure is conceptually similar to factor analysis, and the standardized loadings of the indicators on the latent constructs are similar to factor loadings. Together, the fit index and the pattern of loadings of indicators on constructs test the measurement model. When predicted loadings of indicators on constructs are high and the specified model provides a good representation of the observed data (i.e., good fit index), then the measurement model is demonstrated to be acceptable. If the measurement model has a poor fit, it indicates the assessment of one or more constructs is not valid; either the measurement model needs to be respecified based on available theory or the investigator needs to do more research on assessment of constructs.

Examples of measurement models are presented in Fig. 4. In the model in Fig. 4-I, indicators of parental support and parent-child conflict are presumed to measure two separate constructs; the measurement model is successful and the correlation between the constructs is moderate. In Fig. 4-II, the measurement model includes a construct for peer substance use, specified with three manifest variables, and a construct for adolescent substance use, specified with four manifest variables. The measurement model was found to fit the data well, and therefore was the

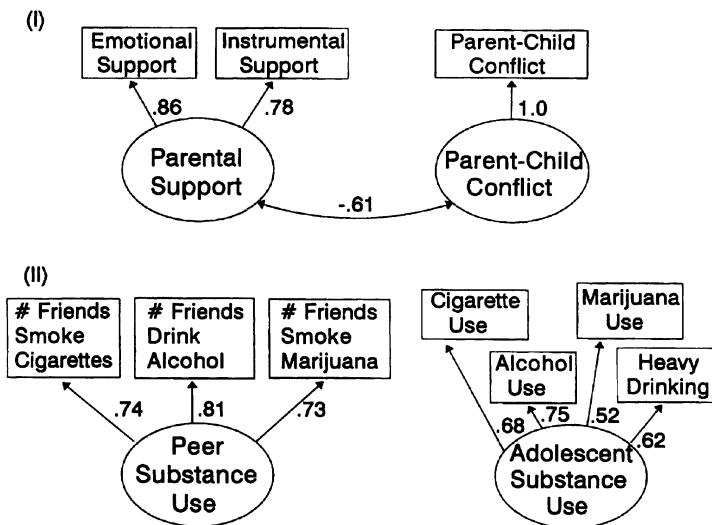


Figure 4. Measurement models: (I) parental support and parent-child conflict; and (II) peer substance use and adolescent substance use.

basis for the subsequent structural model testing the relationship between peer use and adolescent substance use.

How are variables for a measurement model selected? One answer is that indicators are selected by the same process as the predictors, using some combination of theory, clinical experience, and prior research. Theory may suggest that several different types of variables are reflecting a common construct (e.g., Donovan & Jessor, 1985). Clinical experience and pilot research may indicate tests and measures that provide multiple indicators of an interesting construct. Finally, prior research may indicate correlations among related variables that suggest they would be appropriate for a latent construct (e.g., Martin et al., 1994).

Single and Multiple Criterion Constructs

The previous discussion has described models with several predictors and one criterion construct, analyzed to test predictions about the process through which the predictors are related to the outcome. Structural modeling also allows models with multiple criterion constructs, as in MANOVA. For example, Newcomb and Bentler (1988) studied the effects of drug use, social conformity, and social support during adolescence on several different outcomes measured 8 years later, during young adulthood. Outcome constructs included psychosomatic complaints, emotional distress, and problems associated with drug use, health, work, relationships, and the family. In structural modeling the two or more criterion constructs are specified with covariances among them because they are at the same conceptual level (see Figs. 3-V and 3-VI). This procedure has advantages for situations where the researcher has reason to believe that a set of predictors will be related to several

types of outcomes. When it is feasible to collect data for different outcome constructs, the efficiency of a study may be considerably increased.

TESTING THE MODEL

The researcher has developed his or her thinking into a theoretical model that produces predictions about an interesting question, has used the model to design a study, and has finished collecting the data. Now the adequacy of the model can be tested. In this section we outline the analyses and procedures used to test a model. This is the most technical stage of the procedure as it involves programming the analysis, examining relevant parameters that index how well the model fits the data and sometimes making changes to the initial model. We do not give extensive discussion of technical details here. Readers new to the area will likely have the best experience if they can obtain consultation from an experienced researcher or statistician so as to "get the hang" of working with these procedures.

Testing Models in Multiple Regression

The reader is probably familiar with testing predictions about models like those in Fig. 1, using multiple regression analysis. Here the prediction is that a relationship exists and/or that some relationships occur as moderation effects. In this case the multiple regression analysis tests the model by determining whether each effect exists (i.e., is the regression coefficient significantly different from zero) and has the direction predicted by the investigator. A moderation prediction tested by determining whether the cross-product of two predictors is significant when entered with the main effects and whether the interaction has the form predicted by the investigator (e.g., a buffering effect was predicted and an interaction consistent in form with buffering was found). In either case, if the results were significant and consistent with prediction, then the researcher would have evidence that the model was successfully tested. It is desirable to demonstrate that the predicted relationship is significant in a prospective analysis (Fig. 1-II) as well as in cross-sectional models or that a predicted effect remains significant when plausible confounders are included in the regression (Fig. 1-III). However, in any of the four cases one could say that if the predicted relationship was significant, then the researcher's model is supported.

We also note that most of the models in Fig. 3 could be tested with multiple regression analysis. Indeed, for situations where all the variables are manifest variables, the parameters in the model representing relationships between variables (i.e., regression coefficients) would be the same if analyzed in a structural modeling program; here the technique would be termed *path analysis*. For example, in Fig. 3-II, the researcher would perform two regressions, one regression predicting the mediator from the predictor *A* and a second regression predicting the criterion *Y* from the *A* and *M* entered together. This model and more complicated models can be readily analyzed in multiple regression; for example, see Cohen and Cohen (1983) for a discussion of the analytic procedures and Loehlin (1987) for technical discussions of path analysis.

The adequacy of the models in Fig. 3 can be evaluated in multiple regression

but involves more considerations. For one thing, the patterning of results is crucial. In Fig. 3-II, for example, if the investigator predicts a mediated effect, then the path $A \rightarrow M$ and the path $M \rightarrow Y$ must both be significant in order for the researcher to conclude that the model was supported; if either of these paths fails to be significant, then the model is not supported. Conversely, if the investigator predicts there will be a mediated effect and results indicate that only the direct effect is significant, the investigator's thinking is not supported even though significant results were obtained to guide his or her further work. (Note that in this case the zero-order correlation, $A \rightarrow Y$, would be the same in both cases, but the results of a theory-testing procedure would indicate decisively which model was supported: a mediated model or a direct effect model.) In addition, the researcher can determine how much of the variance in the criterion is accounted for by the model, through examining the squared multiple correlation (R^2) for Y . For example, in most cases the researcher would be happier if the model accounted for 10% of the variance in a meaningful criterion than if the model accounted for 1% of the variance in the criterion. Hence, the test of a model is both more demanding and potentially more informative. It requires a clear specification of what model is being tested but provides greater information if the model is correct.

Evaluating Model Adequacy in Structural Modeling Procedures

As we noted previously, structural modeling procedures allow greater flexibility in testing a hypothesis. In addition to allowing a more powerful measurement of constructs and a specific test of the adequacy of the measurement model, structural modeling involves more options for specifying a model than are available in multiple regression and it provides more indicators of the adequacy of the model. These procedures are available in statistical packages such as LISREL (Jöreskog & Sörbom, 1988), EQS (Bentler, 1989), or AMOS (Arbuckle, 1995).

Fit of the Model to the Data

For understanding the parameters available in structural modeling it is important to understand how structural modeling operates. In essence there are five steps. In step 1 the researcher inputs the matrix of n actual covariances* among the variables as measured in the study; these are now termed the *observed* covariances. In step 2, the researcher specifies a model to be analyzed, including specifications of what indicators measure what constructs (i.e., measurement model) and how a construct is related to other constructs (i.e., structural model). In step 3, the analysis creates a statistical model, including specified parameters and parameters estimated from the data. In step 4, the analysis uses this statistical model to compute estimates of the covariances among the measured variables. In step 5, the analysis compares the estimated covariances (from step 4) with the observed covariances (from step 1). Through procedures conceptually familiar to the reader, this comparison produces a χ^2 statistic based on the difference between observed and predicted values, summed over the covariances. This is termed a *fit index*, because

*Correlations (which are really standardized covariances) can also be used as input, but covariances are generally preferred.

it provides a numerical index of discrepancy between the observed values and the model-estimated values; to the extent that there are only small differences between observed and model-estimated values, the χ^2 will be relatively small and the investigator can conclude that the model has good fit. Ideally, the χ^2 test indicates a nonsignificant result, for example, if a χ^2 with 10 *df* ($N = 150$) = 4.80, then this is nonsignificant ($P > 0.10$) and the investigator could say "the statistical model reproduces the observed values well" or more commonly "the model fits the data." To the extent that discrepancies between observed and predicted values were generally substantial, the χ^2 is large and the investigator concludes that this model does not provide a good representation of the data, either because the theory is wrong or the statistical model was incorrectly specified. This procedure is the basis for indexes that test the adequacy of the model, such as the goodness-of-fit index and adjusted goodness-of-fit index (Jöreskog & Sörbom, 1988).

Alternative Fit Indexes

Several fit indexes may be encountered in current literature. The χ^2 statistic is commonly reported as a fit index because it maps directly onto the nature of the structural modeling process, that is, comparing an observed covariance matrix with a predicted covariance matrix. However, the χ^2 index is quite sensitive to sample size, and when samples are even moderate in size, it is typical to find a significant χ^2 test even when other indexes indicate the model is acceptable. Alternative fit indexes have been developed to try to reduce the dependence of fit statistics on sample size. One is the χ^2 -*df* ratio, that is, the actual value of the χ^2 divided by the degrees of freedom in the model; the rule of thumb is that ratio values around 2.0 for large samples provide evidence of reasonable fit. For example, if the χ^2 (100 *df*, $N = 900$) was 180.0, then the χ^2 test would be significant ($P < 0.001$), but the χ^2 -*df* ratio of 1.8 would provide evidence that the investigator had a reasonable model. Another common index is Bentler's comparative fit index (CFI), which reflects the difference between the specified model and a null model; here, CFI values of .90 or greater are taken as evidence of reasonable fit (Bentler, 1989, 1990). Other incremental fit indexes include the Tucker-Lewis index and the normed and nonnormed fit index (Tucker & Lewis, 1973). Detailed discussion is provided in several places (Bollen, 1989; Hu & Bentler, 1995).

With the addition of fit indexes, the evaluation of model adequacy in structural modeling is analogous to that discussed previously for multiple regression. In the case of a successful test, the researcher would find that parameters in the measurement model indicated reasonably high and consistent loadings of indicators on constructs*; predicted relationships between constructs were significant and were in the predicted directions; the χ^2 test (or alternative index) provided evidence of reasonable fit; and the constructs in the model accounted for a respectable amount of variance in the criterion construct. If predicted relationships failed to occur and/or the pattern of relationships among constructs in the model was very different from what was expected, then the researcher has to conclude that his or her theory

*In structural modeling analysis as described here, the analysis is a simultaneous test of the measurement model and the structural model. It is common for the researcher to conduct analyses of the measurement model (i.e., confirmatory analysis) before conducting the final structural modeling analysis.

was not supported. So in this sense, structural modeling provides a straightforward means for testing the fit between the investigator's thinking (as represented in the statistical model) and the real world (as represented by the covariances among measured variables). Structural modeling does provide greater flexibility in specifying either simple or complex models and provides more information about how well the theoretical model fits the empirical data. This makes it feasible to study models with a complex set of direct and indirect effects, which could not be understood simply from multiple regression analysis.

Some comments are in order about how adjustments are made to a structural model. It should be noted that in specifying a model for analysis, the researcher typically defines many parameters to be zero; these may include relationships between constructs that are not predicted to be significant and correlations between error variances, which in many cases are assumed at the outset to be zero (i.e., all the errors of measurement are independent). In current practice, the recommendation is to include in the initial model a minimum number of predicted paths (MacCallum, Roznowski, & Necowitz, 1992). The analysis then provides coefficients and associated *t*-test values indicating the significance of each parameter that was analyzed. It is not uncommon to find that the initial model is unsatisfactory in some ways. Paths indicated as nonsignificant may be dropped from the model by being constrained (i.e., declared to be zero). If the model does not have reasonable fit, adjustments may be made on the basis of modification indexes, which indicate currently constrained parameters that would be significant if included in the model. In some cases, additional relationships between constructs may be included; while failure to find predicted paths may be lethal for an investigator's theory, the finding of nonpredicted but significant effects may provide an opportunity for further research that will help to build further theory if the effects prove replicable. It is also possible that errors of measurement are not totally independent, and the modification indexes will indicate variables whose error variances are correlated. The researcher may then allow what are called correlated error terms to be included in the model. By adding paths or allowing correlated error terms (or *correlated residual terms*), the fit of the model may be improved. Methodologists emphasize that this is not a procedure to be pursued willy nilly, using any means necessary to obtain a better fit index; rather, it is a procedure that should be guided by the investigator's theory. While a theory cannot necessarily address all possible relationships in a complex system, some unpredicted effects may be consistent with the thinking and experience that led to the design of the study, and there is a reasonable basis for adding such effects to the initial model. Possible additions that bear no relation to the investigator's theory are probably best left constrained, perhaps noted as candidates for replication in further studies.

Other Considerations

It is important to recognize that interpreting the adequacy of structural modeling results involves many of the considerations that are relevant for evaluating results from any kind of study. The mere fact that structural modeling has been employed does not relieve the investigator from meeting the burden of proof typically needed on these other issues, and the invocation of structural modeling can do nothing to salvage a study that is weak or faulty on other grounds. The

sample should be reasonably representative of the population to which the investigator wishes to generalize; the list of eligible participants should represent the population and the data should be obtained from a high proportion of those invited to participate. The measures should be ones with demonstrated evidence of validity. The constructs under study should have reasonable variance within the population studied. Sample size is also a consideration, through the ratio of cases to estimated parameters in the model. The recommendation from simulation studies is that a sample size–freely estimated parameters ratio of 5 : 1 is minimally adequate for structural modeling (Bentler, 1989). While this may sound liberal at first hearing, readers should be aware that even a seemingly simple model can have 10–20 parameters, because the model includes indicator loadings, structural paths, and error variances; hence, a sample with 40–50 cases, which is typical for studies designed for *t*-test or analysis of variance (ANOVA) statistics, is none too large for testing such a model. Thus, the researcher should carefully consider sample size issues during the design phase of a study. This does not mean that any structural modeling study must have thousands of subjects, as models have been successfully analyzed with moderate size samples. However, it should be recognized that procedures based on covariance statistics—including multiple regression—require a reasonable number of cases in relation to the number of parameters estimated; and whatever the procedure, smaller samples present limitations to detecting effects that truly exist (Cohen, 1992). These considerations all argue that during the planning phase, the researcher should try to aim for the largest feasible sample rather than the smallest one.

EXAMPLES

In this section, we illustrate the concepts previously discussed with reference to some examples from recent literature in pediatric and clinical child psychology. The aim is not to review the literature in detail but to demonstrate how certain theoretical questions have been tested using the models in Figs. 1–3. The examples were selected to illustrate methodological points, and many studies of comparable interest are not included.

Socioeconomic Disadvantage as a Moderator for Adolescents' Adjustment

An example of a prospective study is an investigation by DuBois, Felner, Meares, and Krier (1994), who used assessments at two time points to examine the effect of socioenvironmental conditions on subsequent adjustment. They noted that previous studies of stress and support had produced mixed results, and they conducted a study with a sample of public school students (mean age, 12.5 years) who were surveyed first at the beginning of the school year and then again 7 months later. Socioeconomic disadvantage was assessed with indexes such as low parental education and student's participation in subsidized lunch programs; predictor variables of stressful life events and perceived social support from family members and school personnel were also assessed. The criterion variables indexing adjustment included psychological distress, drug use, and conduct problems as well as school grades, attendance records, and suspensions.

Prospective multiple regression analyses testing effects of support and stress

by overall number of socioeconomic disadvantages were specified, as in Figs. 1-II and 1-V. The criterion was the time 2 value of the adjustment (Y_2); predictors were the time 1 value for life stress or social support (A) and the time 1 value of adjustment (Y_1); moderators were dummy variables representing single and multiple disadvantage groups (B) compared with the no-disadvantage group. Results indicated significant moderation effects. Among subjects with multiple disadvantages, stress was prospectively related to increased distress and drug use and support from school personnel was inversely related to adolescent problem behaviors, but among subjects from nondisadvantaged homes, these variables had little or no effect on adjustment at time 2. These moderation effects have implications for the design of early interventions, because they show where stress reduction and support-building interventions may have the most impact.

Mediation Model for Parental Socioeconomic Status and Adolescent Risk Behavior

Research has shown children's behavioral problems to be related to some aspects of parents' socioeconomic status (SES), such as education, income, or occupation; but there is less understanding of the process through which these effects occur. A model based on previous research and theory on SES was tested by Wills et al. (1995) with a sample of urban adolescents. Researchers hypothesized that the relationship between parental SES and adolescents' substance (tobacco, alcohol, and marijuana) use would be mediated through variables such as emotional support and academic competence.

A model with multiple mediators as in Fig. 3-IV was tested. Results were as predicted: lower SES of the parents was related to less emotional support in the parent-child relationship and to lower academic competence of the adolescent. These in turn were related to more proximal risk factors for substance use, including deviant peer affiliations. Results were consistent with previous studies, suggesting that economic stress on parents reduced the quality of their marital interaction and their relationships with children; these results "filled in" more details about how these effects are involved in adolescents' behavior problems. The findings had theoretical significance for helping to understand how an abstract demographic variable, SES, is actually related to adolescents' behavior. In addition, it has been noted that cigarette smoking has been increasing particularly among adolescents from blue-collar families, and the findings suggested why these effects are occurring, with implications for smoking prevention programs.

Childhood Chronic Illness and Maternal Mental Health

A group of pediatric psychologists developed a model of the process underlying maternal mental health problems in families with chronically ill children (Lustig, Ireys, Sills & Walsh, 1996), and tested this model with data from a sample of children aged 2 to 12 years old, who were diagnosed with a rheumatological condition, and their mothers. The researchers' model suggested that characteristics such as severity of the illness (both biological and functional impairment), contextual variables such as the age of the child and parent, and availability of supportive services would be related to the outcome.

The test of the model analyzed the direct effects of condition (A) and context

(*B*) characteristics, as illustrated in Fig. 1-III, as well as testing a mediational model hypothesizing that effects of predictor variables on maternal mental health would be partly mediated through the mother's cognitive appraisal of these conditions, as illustrated also in Fig. 3-III where the mediator (*M*) is a score for mother's appraisal. Tests of the mediational model using multiple regression showed that, as predicted, the impact of child's condition on mother's psychological symptomatology was partly mediated through appraisals. The successful test of the authors' theoretical model thus suggests clinical approaches to working with parents of children who have a chronic condition.

Mediational Model for Physical Attractiveness and Psychological Adjustment

A study by Varni and Setoguchi (1996) was designed to clarify previous findings from clinical and nonclinical samples on the relationship between physical attractiveness and psychological adjustment. The participants were adolescents with limb deficiencies who were receiving treatment at a university clinic. The authors hypothesized a model in which the effect of predictor variables was mediated through the adolescent's level of self-esteem, so this is a single-mediator model, as in Fig. 3-II.

Results were obtained using multiple regression analyses with an index of psychological maladjustment (depression and anxiety symptoms) as the criterion. The authors' model was supported; physical appearance had a direct effect to depression–anxiety, but there was also an indirect effect mediated through self-esteem. A measure of general task and social competence also was related to the self-esteem mediator. These findings led to the suggestion that adjustment in adolescents with physical appearance concerns could be improved through counseling interventions that focused on problem solving, expression of feelings, and positive self-talk.

Multiple Criteria Model

A longitudinal data set from the Christchurch Health and Development Study was used to test the pathways from childhood disorder to subsequent outcomes. There were 709 cohort participants, who were assessed in early childhood and at yearly intervals thereafter to age 15 years. The authors questioned whether early conduct problems and attention deficit problems might have different consequences at later years (Fergusson & Horwood, 1995), and a model was specified to test alternative processes suggested by Hinshaw (1992) as explanations for observed correlations between externalizing symptomatology and academic underachievement. A multiple indicator approach was used for the constructs, with early childhood problems indexed through data from multiple sources (parents and teachers), and outcomes at 13–15 years of academic competence and delinquent behaviors indexed through combinations of self-reports, standardized tests, and official record data.

The hypotheses were tested in structural equation modeling. Early conduct problems, attentional problems, and IQ (all assessed at age 8) were specified as exogenous variables. Adolescent academic achievement and delinquent behaviors

(assessed at ages 13–15) were specified as correlated endogenous criterion variables, as in Fig. 3-V. Results showed that the three exogenous variables were substantially correlated, consistent with the typical relationship of conduct problems with other variables. However, even when controlling for the correlations among the predictors, different effects to outcomes were observed. Early conduct problems were related to delinquency in middle adolescence but were not related to academic achievement. In contrast, both IQ and early attention deficit were related to subsequent academic achievement but were not directly related to delinquency. The findings from the model tested by Fergusson and Horwood (1995) serve to rule out some of the explanations discussed by Hinshaw (1992) and provide support for others, drawing attention to the early origins of conduct problems. Moreover, the model indicated the adverse consequences of early attention deficit problems and suggested a need for early intervention for children with this condition.

Differential Pathways for Mediation of Marital Variables

A multiple mediator model was tested by Mann and McKenzie (1996) to represent how marital relationship variables may affect children's oppositional behavior problems. Two parental variables—marital adjustment and open marital conflict—were assessed through standardized scales; several dimensions of parental behavior toward the child were assessed as possible mediators; and children's oppositional behavior was assessed through a checklist that indexed various aspects of noncompliance, resistance, and disobedience.

Analyses were performed in multiple regression with a model specified to test two possible mediators: rejection by mother–father and inept discipline by mother–father (as in Fig. 3-IV). Results indicated no direct effect from parental marital variables to children's behavior problems; rather, the effect of marital variables was totally mediated through the dimensions of rejection and inept discipline. Moreover, differential pathways were noted. The effect of parental marital dissatisfaction on behavior problems was mediated through rejection by father, whereas the effect of marital conflict was mediated through inept discipline by mother. This study lends support to complex models of family systems, and suggests that family therapists may need to give attention to different aspects of the marital relationship in working with fathers and mothers of children with oppositional problems.

SUMMARY

In this chapter we have discussed issues involved in testing theoretical models in child health research. In developing a theoretical model, the researcher will use some combination of clinical experience, prior research and theory, and his or her own unique perspective on a question. Our chapter has presented several analytic models that can be used to test different types of questions and has discussed how such models are analyzed. We have emphasized that models do not need to be complex and arcane in order to be clinically and theoretically meaningful, and a

number of the models we have discussed are relatively simple ones. Yet, we think the concept of formulating a model of the process underlying an outcome and then designing and analyzing data that test this model is shown to be a powerful one.*

We have suggested that there are several advantages to a theory-testing approach. One is that the results can provide more information. If the findings support the researcher's theory, this helps to build confidence in his or her formulation; if the results are not consistent with prediction, they may provide support for an alternative theoretical approach and/or suggest ways in which the design can be modified in subsequent studies. Another advantage of model-testing procedures is flexibility. Structural modeling procedures in particular offer the advantage of specifying a model that is precisely tailored to the researcher's question and making modifications to the model if these are appropriate. A third advantage is efficiency. With analysis of some of the models in Fig. 3, for example, the investigator can design a study using multiple variables to test several questions and can perform an analysis that tests the set of interrelationships among these constructs. For some questions this is preferable to studies that perform piecemeal analyses of one predictor and one criterion at a time, because in such studies it is difficult to understand how a number of conceptually related variables may be linked empirically.

There are aspects of model-testing procedures that some might construe as disadvantages. For one, they are more work because the investigator needs to acquire a working familiarity with relevant theory that could help inform the design of a study; but in going over the examples we have presented of interesting studies from the literature on pediatric and clinical child psychology, the reader may conclude that this is not a particularly bad thing. For another, studies that test processes and pathways involve more risk, because the set of relationships must be consistent with the investigator's prediction and they expose the researcher to the possibility that an alternative model may be shown to be superior. This is a decision that every clinical researcher must make for him- or herself. Finally, the procedures involved in path analysis and structural modeling are more complicated, and some clinical researchers may need a good way of learning the analytic procedures in a user-friendly manner. We suggest that the way to do this is to apprentice with a person who is experienced in structural modeling techniques. Once the novice learns a bit of terminology and sees the procedures in operation with an experienced hand, then the apparent mystery and "mumbo jumbo" of these procedures disappear and the researcher can proceed on his or her own with occasional consultation.

The use of model-testing procedures is opening new perspectives for clinical research. Investigators can test the processes that underlie the effectiveness of a treatment approach and use model testing to test not only the question of "Does it work?" but also the question of how it works. Persons with an interest in clinical

*We should note that some issues are not covered in the present chapter because of space limitations. We are not able to explain procedures for testing alternative models specified for the same variables (nested models). Nor do we explain in detail the procedures for analyzing interaction effects in structural modeling (with multiple-group analysis or "stacked models"). Finally, we do not consider procedures for analyzing highly nonnormal variables. Readers interested in these topics should consult sources such as Bollen (1989), Hoyle (1995), and Jaccard & Wan (1996).

assessment have new approaches for testing the adequacy of different approaches to assessment and diagnosis, not only for defining a disorder itself but also for understanding and measuring the conditions that produce the disorder. Structural modeling procedures provide an efficient way to understand the complexities of adjustment to illness, which typically involve some combination of demographic and environmental characteristics, disease severity and duration issues, and family resources and coping patterns. With univariate analyses it is difficult to get at the “big picture” of adaptation to chronic conditions, but with model-testing procedures a whole set of risk and protective factors can be visualized in one analysis. Longitudinal research methods offer ways to get at unresolved questions about the temporal ordering of predictor and criterion constructs and the emergence of clinical disorder from earlier prodromal characteristics (Hinshaw, 1994; Kazdin, 1995). These are only a few of the questions that are likely to emerge in the coming years for clinical researchers.

APPENDIX

Glossary of Relevant Terms and Concepts

Confounding variable A variable that is related to both the independent and dependent variables in a model and is not a causal variable, and therefore must be controlled for in the analysis.

Covariate Other variables in the model that are not necessarily specified as either independent or dependent variables, for example, demographic variables, but are related to the predictors.

Criterion Outcome or dependent variable.

Dependent variable Criterion or outcome of interest (Y).

Endogenous Variables that are caused or influenced by other variables.

Exogenous Variable that causes and/or precedes endogenous variables.

Independent variable Predictor variable(s) (A).

Latent construct An unmeasurable concept that can be quantified by a group of manifest variables that covary.

Manifest variable A variable is measured with one indicator.

Mediation A mediator variable (M) is an intermediate variable in the causal path from an independent variable to the dependent variable, for example, the effect of variable A on outcome Y is dependent upon the level of variable M .

Model Statistical representation of the relationship between independent and dependent variables.

Model specification Translation of a hypothesis into a mathematical relationship between variables.

Moderation A moderator variable interacts with another independent variable to enhance the predictability of the dependent variable under study.

ACKNOWLEDGMENTS. Equal authorship. This work was supported by a Research Scientist Development Award #K02-DA00252 and grant #R01-DA08880 from the National Institute on Drug Abuse.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.
- Arbuckle, J. L. (1995). *AMOS for Windows. Analysis of moment structures* (Version 3.5). Chicago, IL: Smallwaters.
- Bandura, A. (1969). Social learning theory of identificatory processes. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 213–262). Chicago: Rand McNally.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215.
- Baranowski, T., & Nader, P. R. (1985). Family involvement in health behavior change programs. In D. Turk & R. Kerns (Eds.), *Health, illness, and families* (pp. 81–107). New York: Wiley.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles, CA: BMDP Statistical Software.
- Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, 107, 238–246.
- Biglan, A., Weissman, W., & Severson, H. (1985). Coping with social influences to smoke. In S. Shiffman & T. A. Wills (Eds.), *Coping and substance use* (pp. 95–116). Orlando, FL: Academic Press.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Carver, C. S., Scheier, M. F., & Pozo, C. (1992). Conceptualizing the process of coping with health problems. In H. S. Friedman (Ed.), *Hostility, coping and health* (pp. 167–187). Washington, DC: American Psychological Association.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98, 310–357.
- Cole, R. E., & Reiss, D. (Eds.) (1993). *How do families cope with chronic illness?* Hillsdale, NJ: Lawrence Erlbaum.
- Compas, B. E. (1987). Coping with stress during childhood and adolescence. *Psychological Bulletin*, 101, 393–403.
- Compas, B. E., Malcarne, V. L., & Banez, G. A. (1992). Coping with psychosocial stress: A developmental perspective. In B. N. Carpenter (Ed.), *Personal coping: Theory, research, and application* (pp. 47–63). Westport, CT: Praeger.
- Conger, R. D., Elder, G. H., Jr., Lorenz, F. O., Conger, K. J., Simons, R. L., Whitbeck, L. B., Huck, S., & Melby, J. N. (1990). Linking economic hardship to marital quality and instability. *Journal of Marriage and the Family*, 52, 643–656.
- Doherty, W. J., & Campbell, T. L. (1988). *Families and health*. Newbury Park, CA: Sage.
- Donovan, J. E., & Jessor, R. (1985). Structure of problem behavior in adolescence and young adulthood. *Journal of Consulting and Clinical Psychology*, 53, 890–904.
- DuBois, D. L., Felner, R. D., Meares, H., & Krier, M. (1994). Prospective investigation of the effects of socioeconomic disadvantage, life stress, and social support on early adolescent adjustment. *Journal of Abnormal Psychology*, 103, 511–522.
- Dubow, E. F., & Tisak, J. (1989). The relation between stressful life events and adjustment in elementary school children: The role of social support and social problem-solving skills. *Child Development*, 60, 1412–1423.
- Elder, G. H., Jr., Nguyen, T. V., & Caspi, A. (1985). Linking family hardship to children's lives. *Child Development*, 56, 361–375.
- Fergusson, D. M., & Horwood, L. J. (1995). Early disruptive behavior, IQ, and later school achievement and delinquent behavior. *Journal of Abnormal Child Psychology*, 23, 183–199.
- Flay, B. R., Hu, F. B., Siddiqui, O., Day, L. E., Hedecker, D., Petraitis, J., Richardson, J., & Sussman, S. (1994). Differential influence of parental smoking and friends' smoking on adolescent initiation and escalation of smoking. *Journal of Health and Social Behavior*, 35, 248–265.
- Folkman, S., & Lazarus, R. S. (1985). "If it changes it must be a process": Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology*, 48, 150–170.

- Friedman, L. S., Lichtenstein, E., & Biglan, A. (1985). Smoking onset among teens: An empirical analysis of initial situations. *Addictive Behaviors, 10*, 1–13.
- Friedman, M., & Rosenman, R. H. (1974). *Type A behavior and your heart*. New York: Knopf.
- Garmezy, N., & Masten, A. S. (1991). The protective role of competence indicators in children at risk. In E. M. Cummings, A. L. Greene, & K. H. Karraker (Eds.), *Life-span developmental psychology: Perspectives on stress and coping* (pp. 151–174). Hillsdale, NJ: Erlbaum.
- Gilbert, D. G. (1979). Paradoxical tranquilizing and emotion-reducing effects of nicotine. *Psychological Bulletin, 86*, 643–661.
- Graham, J. W., Marks, G., & Hansen, W. B. (1991). Social influence processes affecting adolescent substance use. *Journal of Applied Psychology, 76*, 291–298.
- Hansen, W. B., & Graham, J. W. (1991). Preventing alcohol, marijuana, and cigarette use among adolescents: Peer pressure resistance training versus establishing conservative norms. *Preventive Medicine, 20*, 414–430.
- Hauser, S. (1991). *Adolescents and their families: Paths of ego development*. New York: Free Press.
- Hauser, S. T., Jacobson, A. M., Lavori, P., Wolsdorf, J. I., Herskowitz, R. D., Milley, J. E., Bliss, R., Wertlieb, D., & Stein, J. (1990). Adherence among children and adolescents with IDDM: Immediate and long-term linkages with the family milieu. *Journal of Pediatric Psychology, 15*, 527–542.
- Hayduk, L. A. (1987). *Structural equation modeling with LISREL*. Baltimore, MD: Johns Hopkins University Press.
- Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111*, 127–155.
- Hinshaw, S. P. (1994). *Attention deficits and hyperactivity in children*. Thousand Oaks, CA: Sage.
- Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology, 65*, 599–610.
- Hoyle, R. L., & Smith, G. T. (1994). Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology, 62*, 492–440.
- Hoyle, R. H. (Ed.). (1995). *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage.
- Hu, L-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.
- Jaccard, J., & Wan, C. K. (1996). *LISREL approaches to interaction effects in multiple regression*. Thousand Oaks, CA: Sage.
- Jessor, R., & Jessor, S. (1977). *Problem behavior and psychosocial development*. New York: Academic Press.
- Jöreskog, K. G., & Sörbom, D. (1988). *LISREL 7: A guide to the program and applications*. Chicago: SPSS.
- Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence* (2nd ed.). Thousand Oaks, CA: Sage.
- Khantzian, E. J. (1990). Self-regulation and self-medication factors in alcoholism and the addictions: Similarities and differences. In M. Galanter (Ed.), *Recent developments in alcoholism* (Vol. 8, pp. 255–271). New York: Plenum Press.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Loehlin, J. C. (1987). *Latent variable models: An introduction to factor, path, and structural analysis*. Hillsdale, NJ: Erlbaum.
- Lustig, J. L., Ireys, H. T., Sills, E. M., & Walsh, B. B. (1996). Mental health of mothers of children with juvenile rheumatoid arthritis: Appraisal as a mediator. *Journal of Pediatric Psychology, 21*, 719–733.
- MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin, 100*, 107–120.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure models. *Psychological Bulletin, 111*, 490–504.
- MacKinnon, D. P. (1994). Analysis of mediating variables in prevention and intervention studies. In L. Beatty & A. Cazares (Eds.), *Scientific methods in prevention research* (pp. 127–153). Washington, DC: U.S. Government Printing Office. (DHHS Publication No. 94-3631).
- MacKinnon, D. P., Johnson, C. A., Pentz, M. A., Dwyer, J. H., Hansen, W. B., Flay, B. R., & Wang, E. Y.-I. (1991). Mediating mechanisms in a school-based drug prevention program. *Health Psychology, 10*, 164–172.
- Mann, B. J., & McKenzie, E. P. (1996). Pathways among marital functioning, parental behaviors, and child behavior problems in school-age boys. *Journal of Clinical Child Psychology, 25*, 183–191.

- Martin, C. S., Earleywine, M., Blackson, T. C., Vanyukov, M. M., Moss, H. M., & Tarter, R. E. (1994). Aggressivity, inattention, hyperactivity, and impulsivity in boys at high and low risk for substance abuse. *Journal of Abnormal Child Psychology, 22*, 177–203.
- Masten, A. S., Morison, P., Pellegrini, D., & Tellegen, A. (1990). Competence under stress: Risk and protective factors. In J. Rolf, A. S. Masten, D. Cicchetti, K. H. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 236–256). New York: Cambridge University Press.
- McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114*, 376–390.
- Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.
- Moffitt, T. E. (1993). The neuropsychology of conduct disorder. *Development and Psychopathology, 5*, 135–151.
- Moffitt, T. E., & Lynam, D. (1994). The neuropsychology of conduct disorder and delinquency: Implications for understanding antisocial behavior. In D. C. Fowles, P. Sutker, & S. H. Goodman (Eds.), *Experimental personality and psychopathology research, 1994: Focus on psychopathy and anti-social personality* (pp. 233–262). New York: Springer.
- Monroe, S. M., & Simons, A. D. (1991). Diathesis–stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin, 110*, 406–425.
- Newcomb, M. D., & Bentler, P. M. (1988). Impact of adolescent drug use and social support on problems of young adults: A longitudinal study. *Journal of Abnormal Psychology, 97*, 64–75.
- Newcomb, M. D., & Felix-Ortiz, M. (1992). Multiple protective and risk factors for drug use and abuse: Cross-sectional and prospective findings. *Journal of Personality and Social Psychology, 63*, 280–296.
- Patterson, G. R., DeBarryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*, 329–335.
- Perkins, K. A. (1995). Individual variability in responses to nicotine. *Behavior Genetics, 25*, 119–132.
- Petraitis, J., Flay, B. R., & Miller, T. Q. (1995). Reviewing theories of adolescent substance use: Organizing pieces in the puzzle. *Psychological Bulletin, 117*, 67–86.
- Ramsey, G. (Ed.). (1989). *The science of family medicine*. New York: Guilford.
- Rosenbaum, M. (Ed.). (1990). *Learned resourcefulness: On coping skills, self-control, and adaptive behavior*. New York: Springer.
- Rutter, M. (1990). Psychosocial resilience and protective mechanisms. In J. Rolf, A. S. Masten, D. Cicchetti, K. H. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in development of psychopathology* (pp. 181–214). New York: Cambridge University Press.
- SAS Institute. (1990). *SAS Technical Report P-200 SAS/STAT Software: CALIS and LOGISTIC Procedures (Release 6.04 edition)*. Cary, NC: Author.
- Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Erlbaum.
- Sher, K. J. (1991). *Children of alcoholics: A critical appraisal of theory and research*. Chicago: University of Chicago Press.
- Stark, K. D., Schmidt, K. L., & Joiner, T. E. (1996). Cognitive triad: Relationship to depressive symptoms, parents' cognitive triad, and perceived parental messages. *Journal of Abnormal Child Psychology, 24*, 615–631.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed). Mahwah, NJ: Erlbaum.
- Takeuchi, D. T., Williams, D. R., & Adair, R. K. (1991). Economic stress in the family and children's emotional and behavioral problems. *Journal of Marriage and the Family, 53*, 1031–1041.
- Tarter, R. E. (1988). Are there inherited behavioral traits that predispose to substance abuse? *Journal of Consulting and Clinical Psychology, 56*, 189–196.
- Thoits, P. A. (1986). Social support as coping assistance. *Journal of Consulting and Clinical Psychology, 54*, 416–423.
- Thompson, R. J., Gil, K. M., Gustafson, K. E., George, L. K., Keith, B. R., Spock, A., & Kinney, T. R. (1994). Stability and change in the psychological adjustment of mothers of children and adolescents with cystic fibrosis and sickle cell disease. *Journal of Pediatric Psychology, 19*, 171–188.
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1–10.
- Turk, D. C., & Kerns, R. D. (1985). The family in health and illness. In D. Turk & R. Kerns (Eds.), *Health, illness, and families* (pp. 1–22). New York: Wiley.
- Varni, J. W., & Setoguchi, Y. (1996). Perceived physical appearance and adjustment of adolescents with

- congenital/acquired limb deficiencies: A path-analytic model. *Journal of Clinical Child Psychology*, 25, 201–208.
- Wallander, J. L., Varni, J. W., Babani, L., Banis, H. T., & Wilcox, K. T. (1989). Family resources as resistance factors for psychological maladjustment in chronically ill and handicapped children. *Journal of Pediatric Psychology*, 14, 157–173.
- Werner, E. E. (1986). Resilient offspring of alcoholics: A longitudinal study from birth to age 18. *Journal of Studies on Alcohol*, 47, 34–40.
- Wills, T. A. (1986). Stress and coping in early adolescence: Relationships to substance use in urban school samples. *Health Psychology*, 5, 503–529.
- Wills, T. A., Blechman, E. A., & McNamara, G. (1996). Family support, coping and competence. In E. M. Hetherington & E. A. Blechman (Eds.), *Stress, coping, and resiliency in children and families* (pp. 107–133). Mahwah, NJ: Erlbaum.
- Wills, T. A., & Cleary, S. D. (1996). How are social support effects mediated: A test for parental support and adolescent substance use. *Journal of Personality and Social Psychology*, 71, 937–952.
- Wills, T. A., & Filer, M. (1996). Stress–coping model of adolescent substance use. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 18, pp. 91–132). New York: Plenum Press.
- Wills, T. A., & Filer, M. (1999). Social networks and social support. In A. Baum & T. Revenson (Eds.), *Handbook of health psychology*. Mahwah, NJ: Erlbaum.
- Wills, T. A., McNamara, G., & Vaccaro, D. (1995). Parental education related to adolescent stress–coping and substance use: Development of a mediational model. *Health Psychology*, 14, 464–478.
- Wills, T. A., Pierce, J. P., & Evans, R. I. (1996). Large-scale environmental risk factors for substance use. *American Behavioral Scientist*, 39, 808–822.
- Wills, T. A., Schreibman, D., Benson, G., & Vaccaro, D. (1994). Impact of parental substance use on adolescents: A test of a mediational model. *Journal of Pediatric Psychology*, 19, 537–555.
- Wills, T. A., & Shiffman, S. (1985). Coping and substance use: A conceptual framework. In S. Shiffman & T. A. Wills (Eds.), *Coping and substance use* (pp. 3–24). Orlando, FL: Academic Press.
- Wills, T. A., Vaccaro, D., & McNamara, G. (1992). The role of life events, family support, and competence in adolescent substance use: A test of vulnerability and protective factors. *American Journal of Community Psychology*, 20, 349–374.
- Windle, M., & Searles, J. S. (1990). Summary, integration, and future directions: Toward a life-span perspective. In M. Windle & J. S. Searles (Eds.), *Children of alcoholics: Critical perspectives* (pp. 217–238). New York, NY: Guilford Press.
- Zucker, R. A. (1994). Pathways to alcohol problems: A developmental account of the evidence for contextual contributions to risk. In R. A. Zucker, G. M. Boyd, & J. Howard (Eds.), *The development of alcohol problems* (pp. 255–289). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

3

Introduction to Individual Growth Curve Analysis

DAVID J. FRANCIS,
CHRISTOPHER SCHATSCHNEIDER,
and COLEEN D. CARLSON

Many questions of interest in child psychiatry, psychology, and neuropsychology are questions about change and its prediction. For instance, do children who suffer closed-head injuries at an early age recover more slowly and show greater long-term deficits than children injured at a later age? Do initial deficits in verbal or spatial functioning lead to deterioration in adaptive functioning after injury? Do characteristics of the child moderate the effectiveness of interventions for attention deficit disorder? To address such questions, researchers frequently employ longitudinal designs that lend themselves naturally to addressing questions about change and its prediction. Unfortunately, some traditional statistical models for longitudinal data are unsatisfactory for characterizing change at the individual level and for studying correlates of change. Consequently, after expending considerable effort to collect longitudinal data, researchers often fail to satisfactorily address questions about change in their analyses. The limitations of many traditional statistical models have led to increased interest in the behavioral sciences in the use of individual growth models for measuring change and examining correlates of change.

An impressive and expanding body of work on growth curve modeling exists in the literature. Numerous papers introducing individual growth curve (IGC) models have been published and convincing arguments for the use of IGC models over traditional methods of analysis of change have been made (Francis, Fletcher, Stuebing, Davidson, & Thompson, 1991; Willett, 1988; Rogosa, Brandt, & Zimowski, 1982). However, these publications vary in their level of technical detail, and as

DAVID J. FRANCIS, CHRISTOPHER SCHATSCHNEIDER, and COLEEN D. CARLSON • Department of Psychology, University of Houston, Houston, Texas 77204-5341.

such, in their usefulness to researchers who are trying to employ the techniques, especially for the first time. In an attempt to "close the gap" between the literature on growth curve modeling and the novice user, this chapter will present a narrow but in-depth illustration of growth curve modeling and its application in neuropsychology in general and developmental neuropsychology in particular. We will begin with a brief review of the limitations of traditional analytic methods, followed by an overview of the conceptual, mathematical, and statistical foundations of individual growth models and the principles that guide their application in neuropsychology. We will pay particular attention to the conceptual meaning of different statistical models that are employed in growth curve modeling by illustrating the different models with a practical example. Finally, we will discuss the major problems associated with the application of individual growth models in neuropsychology.

BACKGROUND

To illustrate the use of IGC models, we selected a subset of 60 children from the Connecticut Longitudinal Study (CLS), a study of the development of reading and academic skills (Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). Three groups of children are represented in this subsample: those who were classified as reading disabled (RD) based on the discrepancy between IQ and achievement at grade 3; low achieving (RD-LA) based on reading achievement less than the 25th percentile at grade 3; and those classified as non-reading impaired (NRI). For the purposes of the present sample, reading classification will be used to predict growth in reading achievement skills across a 9-year time span (grades 1 through 9). Table 1 presents the means and standard deviations for reading achievement and age at the time of testing for each group at each of the nine time points.

Table 1. Means and Standard Deviations
for Reading Achievement Scores at Testing at Each Time Point

Variable ^a	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9
Not reading impaired									
M	462.8	486.6	500.1	508.4	515.6	521.5	526.5	532.3	535.3
SD	(19.1)	(18.5)	(13.7)	(14.2)	(15.1)	(14.2)	(10.6)	(11.1)	(11.5)
N	30	29	30	30	30	30	30	29	30
Reading disabled									
M	427.3	452.2	466.4	480.8	490.7	497.4	504.5	508.4	513.2
SD	(26.2)	(18.6)	(15.6)	(20.5)	(18.8)	(20.1)	(20.4)	(20.3)	(20.8)
N	17	17	17	17	17	17	17	17	17
Low achievement									
M	425.8	456.8	469.2	484.0	490.2	496.1	504.3	508.6	512.5
SD	(15.6)	(8.6)	(7.1)	(10.2)	(9.1)	(9.9)	(9.1)	(9.1)	(9.9)
N	13	13	13	13	13	13	13	13	11

^aM, Mean; SD, standard deviation; N, sample.

Traditional Approaches

Studies of change have traditionally focused on time-related differences in mean performance levels across groups. Changes in mean performance, however, often tell us little about individual change. Instead, examination of mean performance over time assumes that individuals are changing at the mean rate of the group. Consider an artificial example where half of the subjects are growing in a positive direction and half in a negative direction, with considerable intersubject variability in positive and negative growth rates. In this situation, the mean growth rate and consequently the growth in means might be equal to 0. However, a growth rate of 0 may not describe the true growth rate of even one subject in the data set. This example illustrates that mean growth rates potentially convey very little information about growth in the absence of information concerning individual growth and interindividual differences in growth. In fact, in most applied situations, it is far more reasonable to expect significant variability in growth patterns and growth rates than it is to expect all subjects to grow at the same rate and in the same way (e.g., linear, quadratic, exponential, etc.) (Rogosa & Willett, 1985).

In general, there are two basic conceptual approaches to change: the *incremental* view of change, and the *process* view of change. In the incremental view of change, an individual's score at a given time, t , designated Y_{it} , is equal to their score at the prior time point (time $t - 1$) plus some increment. This approach focuses on the stability of individual differences from one point in time to the next. Thus, change is viewed as a series of starts and stops, increments and decrements (Willett, 1988), rather than as a continuous process that takes place throughout the time span under investigation. In Willett's (1988) words, the incremental view of change

is a conceptualization that views individual learning, not as a process of continuous development over time, but as the quantized acquisition of skills, attitudes and beliefs. It is as though the individual is delivered of a quantum of learning in the time period that intervenes between the pre-measure and the post-measure, and that our only concern should be with the size of the acquired "chunk." (p. 347)

Consider, for example, the longitudinal data presented in Table 1. A traditional analysis of this data to address questions regarding the development of reading in these three groups of children might focus on the amount of change that takes place between any two adjacent time points. The difference in reading achievement means between adjacent time points, in essence, becomes the measure of change. Using this approach, Y_{it} is the measure of achievement for person i at time t , and the measure of change for person i from time $t - 1$ to t is simply the difference between two time points, or $D_i = Y_{it} - Y_{it-1}$. This difference can be computed for each person for each pair of intervals (i.e., $T - 1$ difference scores can be calculated for each person measured at a total of T time points).

This difference score approach to measuring change is precisely the incremental approach criticized by Willett (1988). The use of difference scores in analyzing change has received considerable negative attention in the statistical and psychometric literature. Although much of the criticism of difference scores is unwarranted, using difference scores as measures of change is limiting both conceptually and statistically (Willett, 1988). Conceptually, the use of difference scores implies that change occurs in "chunks," and that measuring the size of these chunks is the

objective. Essentially, each new chunk is considered to be present at time t but absent at time $t - 1$. Such an approach also implies explicit interest in the particular occasions of measurement because if a different set of occasions were selected to measure behavior (e.g., observations that were twice as far apart in time), the measure of change would be altered. Thus, the measure of change is confounded with the length of the time interval over which it is being measured, which can complicate the analysis when the length of time between assessments is not constant between individuals.

A second conceptual problem with the incremental approach is that it does not allow the researcher to capitalize on the time-ordered nature of the observations (except between immediately adjacent time points). For example, the addition of 10th grade reading scores to the data presented in Table 1 would provide an additional measure of change (i.e., from grade 9 to grade 10) but would not improve the measurement of reading change from grade 1 to grade 9. Ideally, we would prefer an approach that makes the measurement of change more precise as the number of time points increases.

While there are some significant limitations to the incremental approach to analyzing change, it is incorrect to say that methods do not exist for addressing many of these problems (see Rogosa & Willett, 1983; Willett, 1988; and Francis et al., 1991, for more complete discussions and citations on difference scores). Furthermore, many of these approaches are closely related to the IGC methods reviewed in this chapter. However, they suffer from several important limitations. First, while these approaches can be made to duplicate the individual growth curves approach in certain situations, conducting the analysis is not as straightforward as with IGC methods. Second, and more important, the conditions under which an IGC analysis can be duplicated are limited. Specifically, all individuals must be seen at exactly the same time points and must not have missing data at any time point. In regard to the data in Table 1, this would mean that each subject would need to have been examined in all nine grades, and that the timing of the individual assessments would need to have been identical. When conducting longitudinal research, these two conditions are often impossible to meet. Even in studies where attrition is markedly low, scheduling problems, missed appointments, or delays due to subject mobility can lead to missed assessments and variability in the timing of assessments across individuals. For example, in the example data, the interval between the grades 1 and 2 assessments ranged from 6 to 18 months, with an average of 11.5 months. In a traditional analysis, the variability in the interval would be ignored and treated as if it were of the same magnitude across all individuals (e.g., 12 months each). The extent of bias introduced by failing to meet the assumption of equivalent spacing across subjects will vary from situation to situation. However, if the factors that correlate with change are also the determinants of follow-up intervals (e.g., if low-growth subjects are more likely to change schools and are consequently subject to longer follow-up intervals, or missed assessments), the bias can be substantial.

Individual Growth Curves

An alternative to the traditional focus on mean growth and the incremental approach to change is a process view of change. The process view of change, which

reflects analyses of IGC models, considers a person's score as reflecting an ongoing underlying process that allows for change in the expression of a characteristic (e.g., development or deterioration in behavior). Such a model is more generally characterized as:

$$Y_{it} = \pi_{0i} + \pi_{1i}a_{it} + \pi_{2i}Q^2_{it} + \dots + \pi_{k-1i}a^{k-1}_{it} + E_{it} \quad (1)$$

The subscript i signifies a given individual subject and t represents a given point in time from 1 to k . The π is the individual growth parameters, where the numeric subscript represents the order of the polynomial term (e.g., a 0 subscript indicates a constant term or intercept, 1 a constant rate of change or linear growth rate, 2 the rate of acceleration or quadratic growth rate, etc.). At most, the model can be a polynomial of order $k - 1$ with k time points, although it will usually be a polynomial of substantially lower order. In the study of change, a_{it} is a marker of time. In developmental applications, a_{it} typically reflects the age of subject i at time t , but other equally valid "time" metrics are possible, depending on the specific application. For example, other time metrics might include the number of trials (e.g., in a learning experiment), the number of treatment sessions (e.g., in studying the effects of cognitive rehabilitation), the time since diagnosis (e.g., in studying the response to illness or disease), the time since injury (e.g., in studying recovery from head injury), or the level of drug dosage (e.g., in a dose-response study with individuals receiving multiple doses of the same drug). It is important to point out that the subscript i associated with the growth parameters in Eq. (1) indicates that the parameter values are subject specific (i.e., the growth parameters vary across subjects). Therefore, a growth curve will be estimated for each individual. Last, the term E_{it} represents random error in Y for subject i at time t . In its simplest form, it is assumed that E is normally distributed for all i and t , and that the E s are uncorrelated across subjects at all time points. While it is not necessary to assume that the E s are also uncorrelated across time for a given subject, most specialized statistical software for conducting growth analyses will make this assumption by default.

The goal of the process view of change is to allow change to be described for each individual separately, for groups of individuals collectively, and for the correlates of change to be examined. In fact, it is the ability to study correlates of change more precisely, which gives IGC models one of their greatest advantages over analyses based on difference scores. Let us examine how correlates of change can be studied in an IGC model. Because the parameters of the IGCs describe individual change, correlates of change will relate systematically to variability in the individual growth parameters. Studies examining correlates of change require a second model where the individual parameters, the π_{ki} of Eq. (1), are the dependent variables. Specifically, each π_{ki} is modeled according to Eq. (2):

$$\pi_{ki} = \beta_{k0} + \pi_{k1}X_{k1i} + \beta_{k2}X_{k2i} + \dots + \beta_{kp-1}X_{kp-1i} + R_{ki} \quad (2)$$

where there are $p - 1$ measured characteristics (X_{kp}). The β s represent the effects of the p th characteristic on the k th growth parameter, and R_{ki} is random error. Note that there is no need to assume that all growth parameters are a function of the same measured subject characteristics.

From a psychological perspective, it is the adoption of models (1) and (2) over more traditional models of change and the decisions associated with determining an appropriate form for Eq. (1) that pose the greatest barriers to the analysis of IGCs.

From a statistical standpoint, the critical question becomes how to estimate the parameters of the IGCs most effectively and to appropriately analyze the variability in these estimates [i.e., appropriately estimate the parameters of Eq. (2)]. One approach to estimating IGC model parameters is to use ordinary least-squares (OLS) to estimate Eq. (1) separately for each subject. The OLS estimates from these individual subject analyses of Eq. (1) are then used as the dependent variables in Eq. (2), with the parameters of Eq. (2) (i.e., the β s) typically estimated using weighted least-squares (WLS). There are two major problems with this approach. First, a minimum of two data points is required for any subject whose slope and intercept terms are being estimated, and more are needed if the appropriateness of the IGCs is to be assessed (i.e., a straight line can always be perfectly fit to two data points). Consequently, with only two data points, the subject's growth must follow a straight line (i.e., constant growth), and subjects with fewer data points must be dropped from the analysis. A second disadvantage of the OLS approach is that it yields inappropriate standard errors for the parameter estimates of the individual growth parameters (Bryk & Raudenbush, 1987; Willett, 1988) and frequently inappropriate standard errors for the parameter estimates of Eq. (2). As a result, the search for correlates of change in estimating the parameters of Eq. (2) will be seriously compromised because of inflated type I error rates, or alternatively because of low power.

In contrast, hierarchical linear models, or multilevel models (MLM), represent an alternative to the OLS methodology for analyzing IGCs and correlates of change (Bryk & Raudenbush, 1992). This approach is a development in statistical methodology that has the analysis of change as one of its many applications (Goldstein, 1987) and is applicable wherever data structures are nested. For example, in a study of curriculum-based intervention programs (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998), observations may be made on children but curriculum interventions typically are administered to classrooms. Consequently, the data structure is nested, with children representing the lowest level of the hierarchy and classrooms the second level. The analysis of individual growth represents a special case of the nested data structure, as multiple observations are nested within the individual subject on which those observations are made (Bryk & Raudenbush, 1992). In such analyses, time is considered to be nested within individual.

In analyzing data from nested designs, specifying the appropriate unit of analysis, while critical, is often problematic. The primary difficulty is balancing the lack of independence among observations at lower levels of the hierarchy with the low power for testing hypotheses at the highest level of the hierarchy, where observations are independent. This difficulty translates into the specification of an appropriate error term for statistical comparisons. However, recent developments in statistical methodology accomplish this balance by using advances in maximum-likelihood and generalized least-squares estimation to allow formulation of MLMs and simultaneous estimation of parameters at all levels of the model.

Conceptually, the MLM approach consists of two stages. The first stage consists of a within-subject analysis to estimate the parameters of the individual growth curves. The second stage examines the degree to which individual differences between subjects predict differences in these growth parameters, that is, estimating the parameters of Eq. (2). In actuality, both stages of the analysis occur simultaneously in an MLM approach to the analysis, regardless of the specific criterion

used to optimize parameter estimates (e.g., maximum likelihood or restricted maximum likelihood, to name two typical approaches). However, it is helpful to conceptualize the process as consisting of two stages: a within-subjects multiple regression analysis used to obtain estimates of slope and intercept parameters for each subject, and a between-subject analysis where the intercept and slope parameter estimates become the dependent measures. Growth models are often referred to as *unconditional models* when they involve only the estimation of growth, not the prediction of growth. Models are referred to as *conditional models* when they involve the prediction of growth from subject or group characteristics in Eq. (2).

The availability of two equations [i.e., Eqs. (1) and (2)] involving the growth parameters (i.e., the π_{ki}) creates something of an abundance of riches, statistically speaking. In essence, the two equations give us two estimates of each person's growth parameters. The first estimate comes directly from Eq. (1), and the second estimate comes in the form of a predicted growth parameter based on the estimated values for the β s in Eq. (2) and the individual subject characteristics (i.e., the X_{ij} in Eq. (2)]. Using the estimated β s and the X_{ij} , we can compute an estimate of π_{ki} in much the same way that we compute predicted scores in a regression analysis. Statistically, it would be advantageous if we could combine these two different estimates of π_{ki} into a single optimal estimate, one that makes use of both the relations in Eq. (1) and those in Eq. (2). We compute such an optimal estimate by averaging the two different estimates for a subject. However, to make this average an optimal estimate, we would want to weight the two components that make it up according to their respective precision. The more precise the Eq. (1) estimate is for a person, the more weight we want to give it in computing the average and the less weight we want to give the estimate derived from Eq. (2). It is precisely this differential, precision-based weighting of Eqs. (1) and (2) estimates of the π_{ki} that distinguishes the MLM approach from the OLS and other traditional approaches to estimating IGC models. The MLM approach provides an optimal approach to parameter estimation in so far as it capitalizes on all of the available information for estimating the model parameters and it weights all information proportional to its precision.

Building an Appropriate Unconditional Model

Determining the nature of individual's growth involves both theoretical and statistical considerations (Bryk & Raudenbush, 1992). Developing an unconditional model that appropriately reflects one's data is crucial to modeling individual growth. Misspecification in this model can lead to biased estimates and incorrect standard errors in the conditional models. Thus, an ill-fitting unconditional model is of little use in investigating growth processes. However, the development of growth models also should be guided by parsimony. In the case of building a level 1 model of growth [i.e., Eq. (1)], parsimony dictates the estimation of smaller rather than larger numbers of parameters. Therefore, if growth is adequately determined by a straight line for each subject, then estimating curvature in the growth trajectory (i.e., quadratic growth) would not be necessary. Additionally, if every person's growth rate were estimated well by the same growth rate, then allowing each person to have their own growth rate would not be parsimonious. Thus, in determining the best model with which to estimate growth, while it is important to

"keep it simple," this axiom must be balanced by the determination of the adequacy of the model.

Balancing goodness-of-fit with parsimony is often an iterative process of analyzing models and assessing their fit. To facilitate the following discussion, the iterative process of unconditional model building will be demonstrated using the CLS data presented earlier. In the following examples, various models of growth will be examined and parameters of growth will be examined in terms of both their average values and their variances and covariances. Finally, an assessment of model fit will be performed. Due to the numerous methods of assessing fit and the space limitations of this chapter, we will concentrate on only one method-of-fit inspection, namely, the examination of residuals. See Bryk and Raudenbush (1992) for a thorough treatment of model fit.

In growth curve modeling, the age of the child is often employed as a level 1 predictor of growth. When employing a particular growth curve model, the predicted values of the dependent variable can be computed and compared to a child's actual score on the dependent variable. The difference between these two values gives an estimate of the residuals in Eq. (1) (E_{it}). Expressing these residuals relative to their estimated standard errors gives standardized residuals. Under standard modeling assumptions, these standardized residuals will be distributed normally with a variance of 1.0 and mean of zero for any given value of a_{it} . Thus, a plot of the Eq. (1) standardized residuals against a_{it} will show no relationship between the residuals and age. Moreover, the residuals should fall symmetrically about zero with relatively constant variance. The discovery of age-related variation in the scatter plot of standardized residuals is an indication of poor model fit. Specially, such a pattern indicates that the model has not fully captured age-related variation in the outcome measure, and that the model in Eq. (1) needs to be modified. Nonnormal distribution of the residuals and/or nonconstant variance in the residuals with respect to age indicate that the statistical assumptions underlying the model may not be met. Such problems raise concerns about the estimation of standard errors for model parameters in Eqs. (1) and (2), and consequently tests of significance and hypothesis tests involving model parameters may be inaccurate. Estimation of model parameters under alternative modeling assumptions is beyond the scope of this chapter. Readers may wish to consult the literature on generalized estimating equations (Zeger, Liang, & Albert, 1988) and generalized linear mixed effects models (McCullagh & Nelder, 1989).

Building an appropriate model of growth is an iterative process that requires the examination of various growth models. Not only must the nature of growth be determined (e.g., linear, curvilinear, etc.), but whether each parameter should be treated as fixed or random also needs to be specified. In other words, if a significant amount of variance around the mean value of the parameter exists, then there is evidence that the parameter should be allowed to vary for each individual, that is, to be "random." If significant variance does not exist, then that parameter is typically constrained to be equal across individuals, that is, to be "fixed." For example, a fixed linear slope effect would imply that all individuals are growing at the same rate, whereas a random slope effect would imply growth rates that differ across subjects.

The terminology of fixed and random effects often causes some confusion for students. Some of this confusion stems from the use of these terms in the literature

and the labeling of output used in some statistical software. Students may find it helpful to keep in mind that if a parameter is constrained to have the same value for all subjects, then we need only estimate that single value. However, if a parameter is allowed to vary across subjects, then we will want to estimate both the average value of the parameter and the variance in the values of the parameter across subjects. The average value of the random effect parameter is also referred to as the parameter's "fixed effect," which seems to be the source of confusion for students.

Before building an unconditional model, the investigator must decide on a suitable representation of the "time" metric. The intercept parameter in Eq. (1) must be interpreted relative to the time metric represented by a_{it} . Specifically, the intercept represents expected performance when time equals zero. Oftentimes, a value of zero for the time metric employed (e.g., age) will be meaningless. Therefore, it is recommended that the metric be changed such that the intercept value reflects performance at a meaningful value of the time variable. This process of specifying the value at which time is zero is referred to as "centering the time variable." In the illustrative CLS data, the time metric has been defined as the child's age in years. Furthermore, age was centered such that the intercept value represents a child's expected performance at 14 years of age, that is, age was centered at 14 years.

One of the simplest models that can be estimated is called a random intercept model, or random effects analysis of variance (ANOVA) such that:

$$Y_{it} = \pi_{0i} + E_{it} \quad (3)$$

where:

$$\pi_{0i} = \beta_{00} + R_{0i} \quad (4)$$

Notice that Eq. (3) is a simple version of Eq. (1), and that in Eq. (3), no growth rate parameter is estimated. As such, this model implies that individuals vary only on their mean performance level and that their performance is constant at all time points except for error. A strictly random intercept model rarely provides adequate fit to developmental data. However, in building an unconditional model, this model can be tested prior to the addition of growth parameters.

Table 2 presents the estimates and statistical tests of the fixed and random effects in the model. A significant fixed effect indicates that the average parameter estimate is significantly different from zero, while a significant random effect indicates that there is sufficient reliable variance around the mean to allow the estimation of separate parameters for each individual. In this model, the statistical

Table 2. Estimates for the Random Intercepts-No Slope Model

Fixed effect	Coefficient	Standard error	t-Ratio	df
Intercept	496.1	2.49	199.22	59
Random effect	Variance component	Standard error	Z	P value
Intercept	288.38	68.61	4.20	.0001
Residual	746.93	48.41	15.43	.0001

tests for both the fixed and random intercept effects are significant, indicating that the average of the student-level average reading achievement scores is different from zero ($t = 199.22$; $df = 59$), and that there is a significant amount of variance in individual's average reading scores around this mean ($z = 4.20$, $P \leq 0.001$).

The next step in model building would be the addition of a fixed linear slope effect. The addition of a fixed linear growth term to the above model produces a random intercepts–fixed slope model, or a random effects analysis of covariance (ANCOVA) such that:

$$Y_{it} = \pi_{0i} + \pi_{1i}a_{it} + E_{it} \quad (5)$$

where:

$$\pi_{0i} = \beta_{00} + R_{0i} \text{ and} \quad (6)$$

$$\pi_{1i} = \beta_{10} \quad (7)$$

The $\pi_{1i}a_{it}$ term in Eq. (5) represents the addition of age, or growth, to the model. Thus, while each individual's intercept value is allowed to differ, all individuals are forced to have the same growth rate (or slope). Fixing the coefficient π_{1i} to be equal across individuals [Eq. (7)], forces everyone to have the same slope or growth rate. Table 3 presents the statistical tests of the fixed effects for the intercept and slope estimates, as well as the test of the random intercept effect. The tests of the fixed intercept and slope effects are significant ($t = 197.02$, $df = 59$; and $t = 48.77$, $df = 475$, respectively), indicating that these parameters are significantly different from zero. The test of the random effect of the intercept is also significant, indicating that there is sufficient reliable variance around the mean intercept value to allow the estimation of a different intercept for each person ($z = 5.25$, $P \leq 0.001$). The predicted growth curves for each child appear in Fig. 1a. In this figure, all children are growing at the same rate, but children vary in their estimate of reading achievement at age 14 (in their intercepts). To further determine the adequacy of this model, a scatter plot of the raw residuals by age is presented in Fig. 1b and clearly shows that the residuals are still related to age. Moreover, the pattern of the plotted residuals indicates that this model overpredicts reading achievement scores at both younger and older ages (evidenced by residuals less than zero), while it underpredicts scores in the middle age range (evidenced by residuals greater than zero). Therefore, while the parameters in this model significantly contribute to model fit, they do not provide an adequate representation of the data.

In order to capture the nature of growth in reading achievement, the next step

Table 3. Estimates for the Random Intercepts–Fixed Slope Model

Fixed effect	Coefficient	Standard error	t-Ratio	df
Intercept	523.93	2.66	199.02	59
Slope	9.11	0.19	48.77	59
Random effect	Variance component	Standard error	Z	P value
Intercept	390.78	74.48	5.25	.0001
Residual	124.4	8.07	15.41	.0001

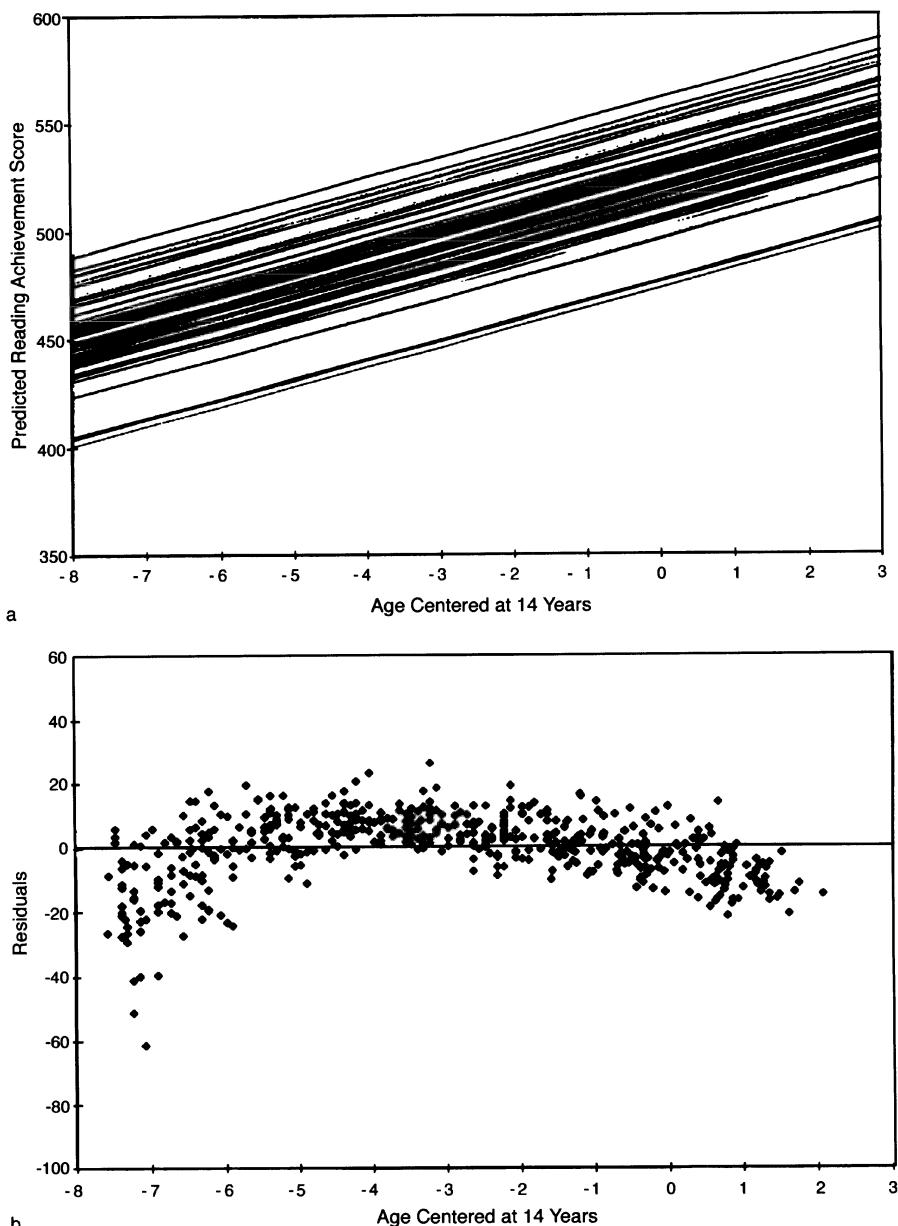


Figure 1. (a) Random intercepts, fixed slope model: Predicted values over time. (b) Random intercepts, fixed slope model: Residuals.

in model building is to allow the slope parameter to be random. This model is called a random intercepts-random slopes model, such that:

$$Y_{it} = \pi_{0i} + \pi_{1i}a_{it} + E_{it} \quad (8)$$

where:

$$\pi_{0i} = \beta_{00} + R_{0i} \quad (9)$$

$$\pi_{1i} = \beta_{10} + R_{1i} \quad (10)$$

The addition of R_{1i} to Eq. (10) allows for the estimation of separate growth curves for each child. As shown in Table 4, the tests of the fixed effects indicate that the average slope and intercept values are significantly different from zero ($t = 217.02$, $df = 59$; and $t = 33.68$, $df = 59$, respectively), while those for the random effects indicate that both of these parameters have enough variance to be estimated separately for each child ($z = 5.0$, $P \leq 0.001$, and $z = 3.19$, $P \leq 0.001$, respectively). Figure 2a illustrates how freeing the slope parameter allows the estimates of growth to differ such that some children show faster growth than others. Examination of the age by residual scatter plot (Fig. 2b) for this model indicates that a relationship between age and the standardized residuals still exists. Comparison of the plots in Figs. 1b and 2b suggests that the present model slightly improves estimation at later ages, but does not change prediction in the early and middle age ranges. However, the pattern of the residuals suggests that allowing for curvature (i.e., a quadratic term) in the growth model may provide a better fit to the data.

The next step is to add a fixed quadratic growth parameter to the model, which is accomplished by squaring the age value and adding this additional variable to the model. Building on Eq. (8), the formula for this model is:

$$Y_{it} = \pi_{0i} + \pi_{1i}a_{it} + \pi_{2i}a_{it}^2 + E_{it} \quad (11)$$

where:

$$\pi_{0i} = \beta_{00} + R_{0i} \quad (12)$$

$$\pi_{1i} = \beta_{10} + R_{1i} \quad (13)$$

$$\pi_{2i} = \beta_{20} \quad (14)$$

The $\pi_{2i}a_{it}^2$ term [Eq. (11)] represents the quadratic term, which allows for curvature in the growth trajectories representing change in reading achievement and is equal across all individuals [Eq. (14)]. As can be seen in Table 5, the addition

Table 4. Estimates for the Random Intercepts–Random Slopes Model

Fixed effect	Coefficient	Standard error	t-Ratio	df
Intercept	523.97	2.41	217.02	59
Slope	9.11	0.27	33.68	59
Random effect	Variance component	Standard error	Z	P value
Intercept	319.87	64.00	5.00	.0001
Slope	2.62	0.82	3.19	.0014
Residual	105.32	7.30	14.45	.0001

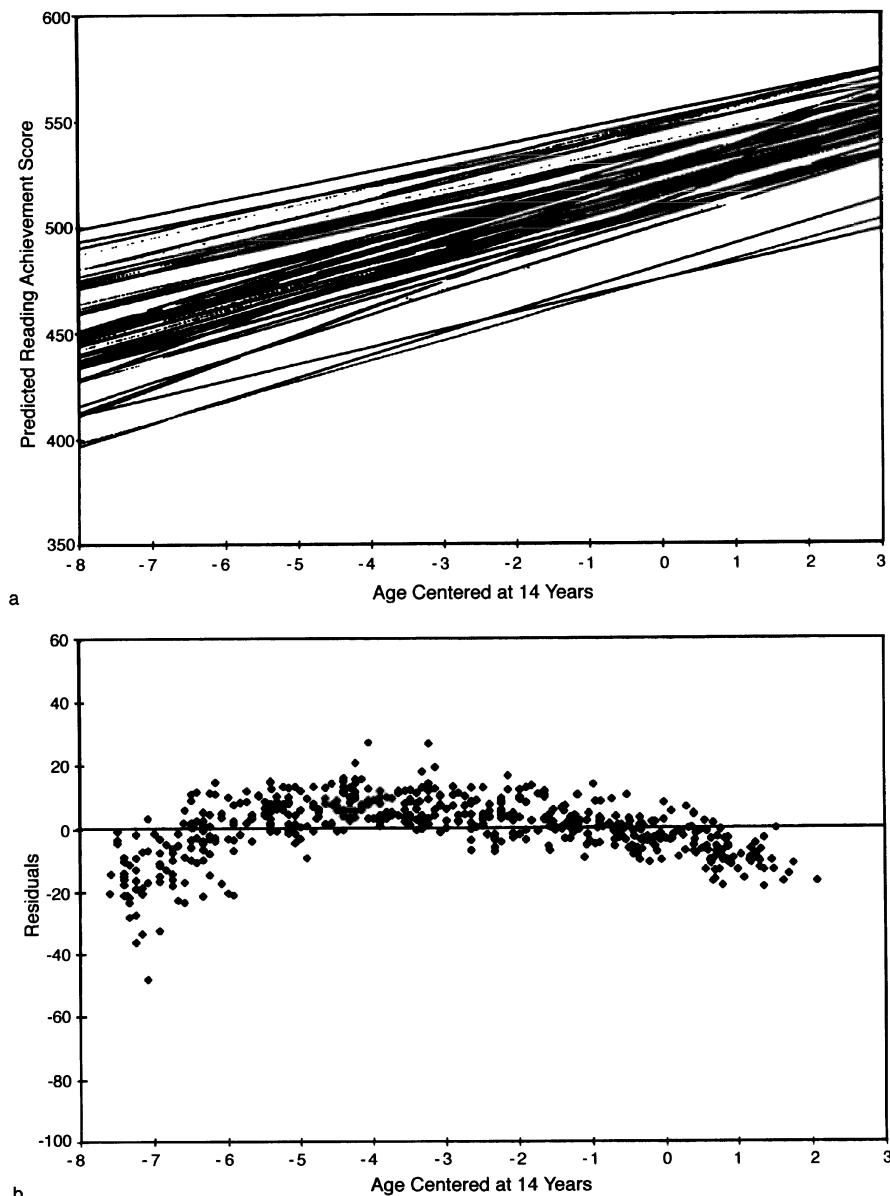


Figure 2. (a) Random intercepts, random slopes model: Predicted values over time. (b) Random intercepts, random slopes model: Residuals.

Table 5. Estimates for the Random Intercepts–Random Slopes–Fixed Quadratic Model

Fixed effect	Coefficient	Standard error	t-Ratio	df
Intercept	520.86	2.30	226.64	59
Slope	2.12	0.41	5.14	59
Quadratic	-1.15	0.05	-22.86	59
Random effect	Variance component	Standard error	Z	P value
Intercept	302.96	58.09	5.22	.0001
Slope	3.81	0.85	4.48	.0001
Residual	47.05	3.26	14.41	.0001

of a fixed quadratic effect contributes significantly to the model ($t = -22.86$, $df = 59$). Visual inspection of the estimated growth curves in Fig. 3a also supports the importance of allowing for curvature in the growth trajectories for reading achievement. This figure clearly illustrates how children's growth rates slow over this time period. Examination of the residual by age scatter plot in Fig. 3b indicates a significant improvement in model fit to the data with the inclusion of a quadratic term. Estimation at all age ranges seems to be improved in this model (evidenced by the greater number of residuals surrounding the zero line).

A further step in the model building for the present example examines the benefit of allowing the quadratic term to be random, i.e., to vary across individuals. The final model is a random intercepts–random slopes–random quadratic model, such that:

$$Y_{it} = \pi_{0i} + \pi_{1i}a_{it} + \pi_{2i}a_{it}^2 + E_{it} \quad (15)$$

where:

$$\pi_{0i} = \beta_{00} + R_{0i} \quad (16)$$

$$\pi_{1i} = \beta_{10} + R_{1i} \quad (17)$$

$$\pi_{2i} = \beta_{20} + R_{2i} \quad (18)$$

The addition of R_{2i} in Eq. (18) allows for the estimation of separate deceleration parameters for each child. Table 6 presents the tests of the fixed and random effects. Results indicate that allowing the quadratic term to be random significantly contributes to model fit ($z = 3.24$, $P \leq 0.001$). Figure 4a also highlights the benefit of allowing the quadratic term to be random, as it illustrates how the curvature of the growth trajectories varies across children. Furthermore, in Fig. 4b, a slight tightening of the residuals around the zero line can also be seen, indicating that a random quadratic term improves model fit to the data.

At this point, one could examine the addition of cubic and higher-order polynomial terms to the model. In the present analyses, there was no theoretical basis for such an addition, and thus, the random intercept–random slope–random quadratic model was determined to be the one that best represented the data. Examining the estimates for this model (Table 6) indicates an estimated average intercept value of 520.86, an estimated average slope of 2.12, and an estimated

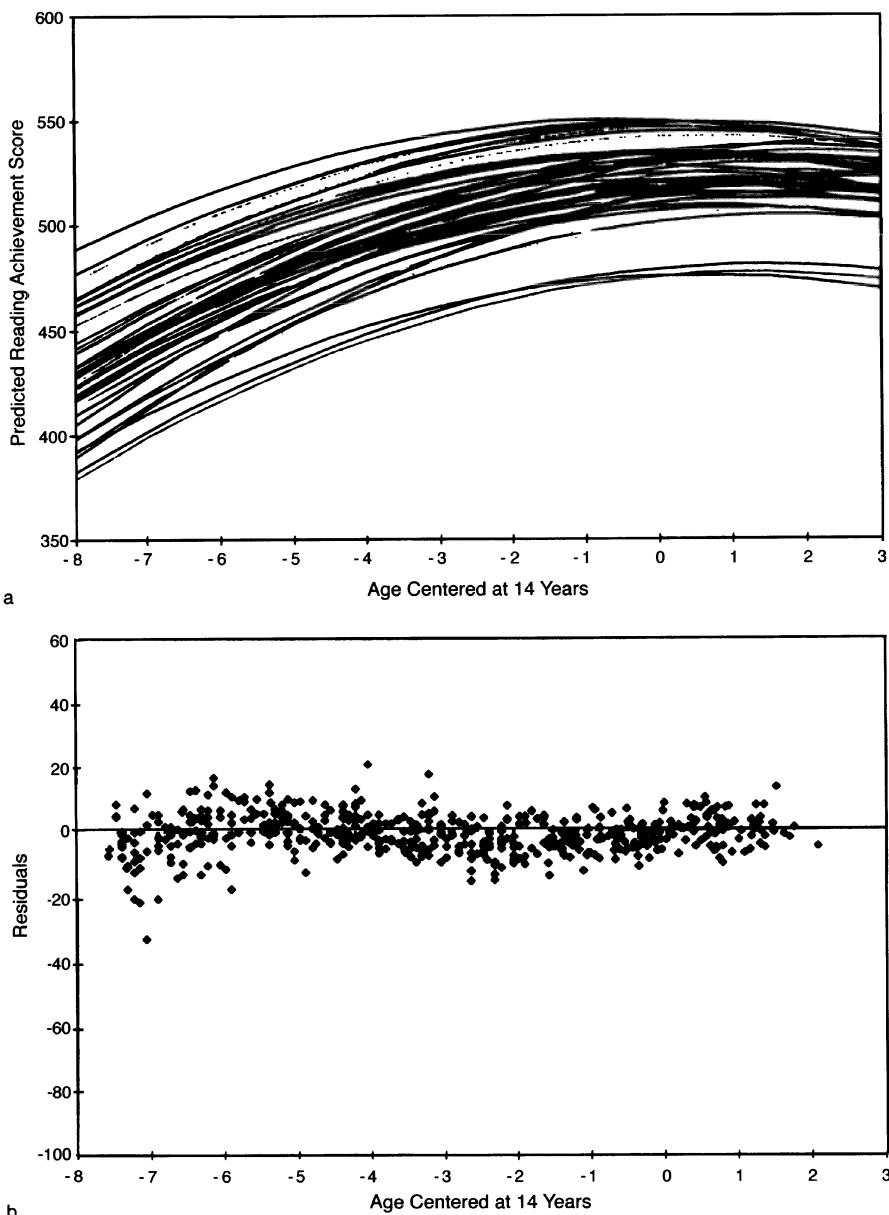


Figure 3. (a) Random intercepts, random slopes, fixed quadratic model: Predicted values over time.
(b) Random intercepts, random slopes, fixed quadratic model: Residuals.

Table 6. Estimates for the Random Intercepts–Random Slopes–Random Quadratic Model

Fixed effect	Coefficient	Standard error	t-Ratio	<i>df</i>
Intercept	520.86	2.30	226.64	59
Slope	2.12	0.41	5.14	59
Quadratic	-1.15	0.05	-22.86	59
Random effect	Variance component	Standard error	Z	P value
Intercept	312.36	59.56	5.24	.0001
Slope	1.81	1.36	1.33	.1827
Quadratic	0.19	0.06	3.24	.0012
Residual	38.55	2.88	13.39	.0001

average rate of deceleration of -1.15. Thus, the estimated average reading score at age 14 is 520.86, and children's reading scores are improving at a rate of 2.12 units per year at that age. [Note, in a straight-line growth model like Eq. (5) or (8), the slope indicates the rate of change in performance at any age. However, in a quadratic growth model like Eq. (11) or (15), the rate of growth, i.e., the slope parameter, varies with age. In this case, the slope parameter for an individual indicates the rate of growth for that individual at that point where age is zero.] Finally, the quadratic parameter indicates that the rate of growth in reading achievement is estimated to be slowing at an average rate of 2.30 (2×1.15) units per year. In the current model, the rate of reduction in the rate of growth differs across subjects, but is constant throughout the trajectory of a given child. In other analyses of the CLS, we have used a model where the rate of deceleration is constant up to a point in development for a given child, at which point the rate of development and the rate of deceleration go to zero. This specification leads to a growth trajectory that follows quadratic growth to a plateau, and then remains stable at that plateau value. See Francis, Shaywitz, Stuebing, Fletcher, and Shaywitz (1996) for more detailed discussion of this alternative model and how it relates to the quadratic model of Eq. (15).

Conditional Model

Once an unconditional model has been chosen, the next step is to add the child-level predictors to the model. Conceptually, this step is similar to building a multiple regression model in which the parameters of the unconditional model become the dependent variables to be predicted by a set of independent variables. In our example, group membership (NRI, RD, LA) is the independent variable. From the unconditional model, we have three parameters to be predicted: the intercept, which represents the mean reading achievement of children at age 14; the slope, which is the average instantaneous growth rate at age 14; and the quadratic term, which represents the average rate of change in growth rates each year over the 9 years of the study. As in multiple regression, a coding scheme must be employed to represent the three levels of group measurement (Cohen & Cohen, 1983). The decision regarding how to best represent group membership should be guided by the theoretical questions of interest. In this example, we were most interested in whether the RD and the LA groups differed in the development of reading achieve-

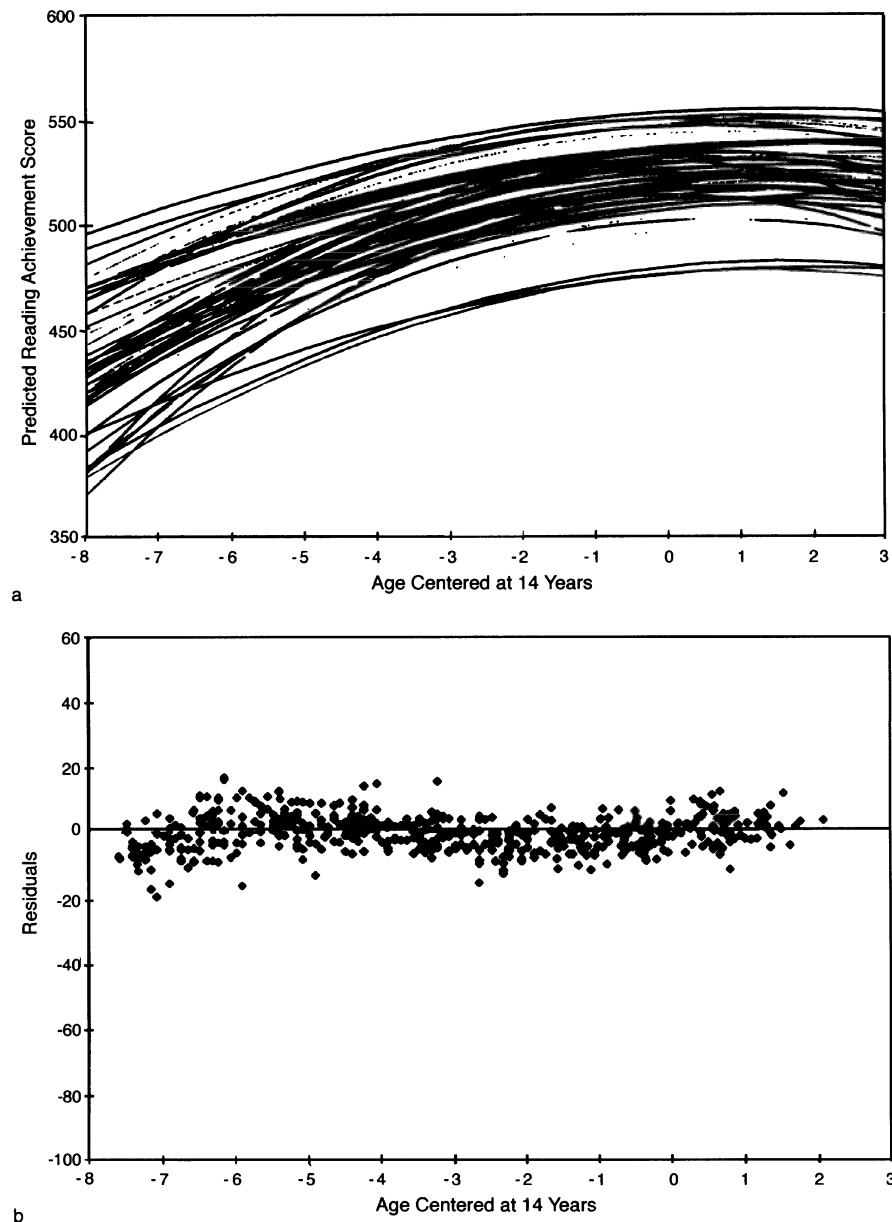


Figure 4. (a) Random intercepts, random slopes, random quadratic model: Predicted values over time.
(b) Random intercepts, random slopes, random quadratic model: Residuals.

ment. Group membership was dummy-coded so that the direct comparison of LA and RD and of LA and NRI was possible. This was accomplished by setting the LA group value equal to 0.

The results of the conditional model can be seen in Table 7. Because of our coding scheme, the intercept value in the fixed effects table (509.48) represents the predicted mean reading achievement of the LA group at 14 years of age. The next parameter (23.12) represents how many units higher the NRI group mean is over the LA group at 14 years of age. The *t*-test associated with this parameter indicates that the NRI group reading achievement scores are significantly higher than the LA groups' mean at 14 years old. The third fixed effect coefficient in the table (-.06), which is nonsignificant, represents the mean difference between the LA and RD groups at 14 years of age.

The coefficients for the slope and the quadratic term can be interpreted in the same fashion. There are no significant differences between the LA and the NRI or RD groups in either the linear or quadratic growth parameters. However, since the nature of growth in this model is determined by both the linear and quadratic terms simultaneously, it can sometimes become clumsy to address significant differences in the linear and quadratic terms separately. Often it is useful to take the parameters from the conditional model and calculate the equations that describe growth in each group. For example, the equation that describes growth for the LA group is:

$$Y = 509.48 + 2.47(\text{age}) - 1.27(\text{age}^2)$$

The equations for the two other groups can also be calculated from the parameters in Table 7. Similarly, the growth model for the NRI group is:

$$\begin{aligned} Y &= (509.48 + 23.12) + (2.47 - .54)(\text{age}) + (-1.27 + .24)(\text{age}^2) \\ Y &= 532.6 + 1.93(\text{age}) - 1.03(\text{age}^2) \end{aligned}$$

and for the RD group is:

$$\begin{aligned} Y &= (509.48 - .06) + (2.47 + .01)(\text{age}) + (-1.27 + .03)(\text{age}^2) \\ Y &= 509.42 + 2.48(\text{age}) - 1.24(\text{age}^2) \end{aligned}$$

Table 7. Estimates for the Conditional Model

Fixed effect ^a	Coefficient	Standard error	<i>t</i> -Ratio	<i>df</i>
Intercept	509.48	3.86	132.00	57
NRI	23.12	4.62	5.00	356
RD	-0.06	5.12	-0.01	356
Slope	2.47	0.41	5.14	57
NRI	-0.54	0.88	-0.61	356
RD	0.01	0.97	0.01	356
Quadratic	-1.27	0.15	-8.21	57
NRI	0.24	0.19	1.32	356
RD	0.03	0.21	0.70	356
Random effect	Variance component	Standard error	<i>Z</i>	<i>P</i> value
Intercept	182.45	36.20	5.04	.0001
Slope	1.99	1.40	1.41	.1572
Quadratic	0.19	0.06	3.18	.0015
Residual	38.53	2.88	13.39	.0001

^aNRI, non-reading impaired; RD, reading disabled.

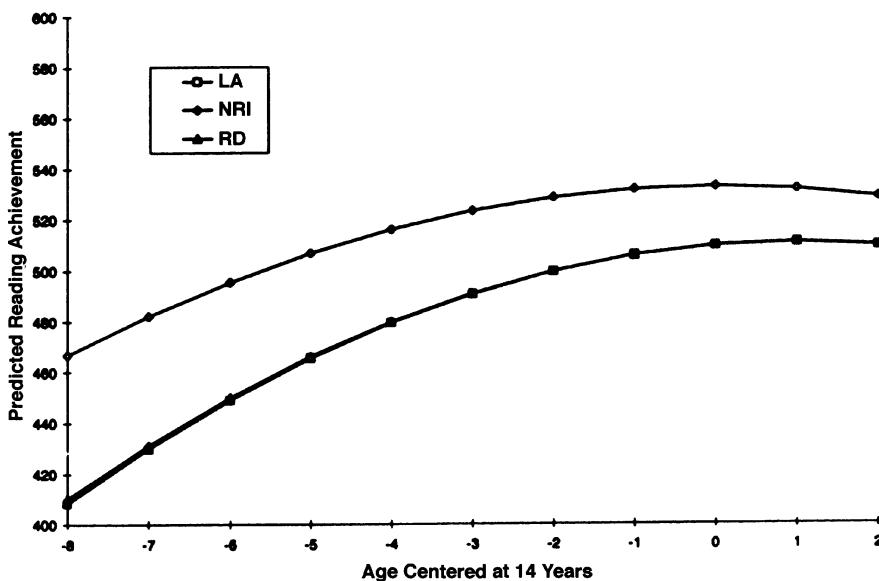


Figure 5. Predicted growth in decoding for low achieving, reading disabled, and children without reading impairment.

The predicted growth curves are shown in Fig. 5. This graph clearly depicts the higher overall performance of the NRI group, as well as the similarity of all three groups in terms of the nature of growth over time. Thus, whether a child is classified as reading disabled or low achieving, their developmental trajectories in reading achievement are almost identical. Furthermore, "catching up" to normal readers on average does not appear to occur for the reading impaired groups.

CONCLUSION

There are several advantages to the MLM approach to the study of change over traditional methodologies such as residualized change score analysis and repeated measures ANOVA. Specifically, these latter approaches look only at mean or incremental change, whereas MLM allows examination of the form of individual growth, its correlates, and mean growth. MLM has additional advantages over alternative methodologies for IGC analysis (Bryk & Raudenbush, 1987). Specifically, MLM offers greater power to detect effects of between-subjects characteristics on growth parameters due to greater precision in estimation of the within-subjects growth parameters. MLM also allows separate estimation of variance in slopes and intercepts due to sampling error and variance due to subject differences in true slopes and intercepts. Consequently, it is possible to form estimates of the reliability of slope and intercept estimates that addresses the potential for the outcome measures employed to demonstrate differences in change (Bryk & Raudenbush, 1987). Another advantage of MLM to the study of change is that the analysis makes use of

whatever within-subjects data are available for a given subject, that is, it is not necessary to eliminate subjects from the analysis because they failed to make an appointment (Bryk & Raudenbush, 1987).

This last point is a considerable advantage for MLM in many neuropsychological research settings. The use of MLM allows us to get the most out of the available data for studying change. This is not an excuse for sloppy design or data collection, but a recognition of the difficulty of getting complete data sets on individuals over a protracted time frame, particularly when those individuals suffer a serious behavioral disorder that may affect compliance. Any complete analysis of such a data set would include a thorough investigation of the missing data, the factors that affect whether the data are observed or missing, and consideration for any bias that might result from nonrandom missing data.

It is instructive to review how MLM accomplishes this maximization of available information. In an MLM analysis each subject's estimate is a weighted average of an estimate based on that subject's individual data and an estimate based on the conditional average of all subjects, conditional on the subject characteristics entered at the "second stage" of the analysis [i.e., in Eq. (2)]. The weight applied to the individual data is directly proportional to the reliability of that data, which is usually a function of the number of data points, the variability in the time points, and the degree to which the outcome measure conforms to the proposed individual growth model. The weight applied to the average estimate is inversely proportional to the reliability of the individual data. Thus, as precision in the individual estimate increases, the individual estimate receives greater weight. As precision in the individual estimate falls, the conditional group mean estimate receives greater weight. As Bryk and Raudenbush (1987) conclude:

In short, (M)LM capitalizes on any strength in the data. If the individual growth trajectory estimates are reliable then (M)LM weights them heavily. If the latter estimates are not reliable, the model substitutes values from mean growth trajectories that are conditional on available background information. (p. 155)

This weighted estimate of the individual growth parameter has been shown theoretically and empirically to have smaller mean squared error than either of the two estimates that go into its derivation, that is, the estimate based only on the individual's repeat observations and the estimate based on the group average trajectory (Efron & Morris, 1979; Morris, 1983). Consequently, the MLM approach offers a presently optimal analytic strategy for the study of change.

One final comment concerns the measures of cognitive performance to be used in growth analysis. Because change has not been the predominant focus of social science research, the measures employed in most psychological research are designed to discriminate between subjects at a fixed point in time. Moreover, transformations are frequently applied to test measures in order to stabilize the variability in test scores over time. The use of tests designed to yield scores with equal variance over time virtually guarantees that information regarding heterogeneity in growth will be lost (Rogosa et al., 1982). Rather, it is important that the study of growth incorporate measures capable of demonstrating growth and capable of demonstrating individual differences in growth when such differences exist. For example, in their traditional form, standardized intelligence tests such as the Wechsler Intelligence Scales for Children—Revised (WISC-R) and Stanford-Binet are problematic

for use in the study of growth for the very reason that the standard scores are subjected to different transformations at specific ages in order to maintain constant mean and variance in the population at all ages. Loss of standardized instruments from the study of change would be unfortunate, because these measures represent some of the most carefully designed and psychometrically sound instruments available to neuropsychologists. However, it is unnecessary to sacrifice these measures simply because they yield standardized scores under normal reporting procedures. The design of many of these instruments and their method of administration actually permits their use in the study of change, provided one does not transform the obtained raw score point totals to standard scores. For tests of this nature, scoring remains constant regardless of the age of the subject, and the same items are administered throughout the age range of the test when the same form is used (or comparable items when alternate forms are used). Consequently, higher raw scores over time indicate improved test performance in the sense that more items were answered correctly. Whether one's relative position remains stable (standard score consistency) need not be of concern, as this information will be reflected in the individual growth curves. The most serious drawback to the use of raw scores instead of traditional standard scores from these and other tests is the arbitrariness of the raw score metric in which most of the scales are based. Consequently, changes in the precise form of the growth curves as a function of the base level of performance may reflect problems in test scaling across the range of available scores as much or more than the shape of some "true" growth function, that is, the shape of the curve describing growth in the construct presumed to be measured by the test. However, this limitation of the raw score metric does not prevent comparison of growth curves for individuals performing at comparable overall levels. That is, one could not use this limitation to explain why two subjects with mean performance levels of 50 have slopes that differ in direction or even magnitude. Thus, this limitation to the use of raw scores from available, psychometrically sound instruments seems minor in comparison to the lost information when standardized scores are used in the analysis of change, and in comparison to the cost and difficulty of developing an entirely new set of instruments for use in the measurement of change. If alternate forms of a test are used, then it is likely that some scaling work would need to be done to equate scores on the alternate forms before a meaningful growth curve analysis could be done.

In the case of instruments that can only be used in their standardized metric, an individual growth analysis is still possible. However, in this case it is only possible to distinguish between patterns of growth and not differences in true "rates" of change. That is, we can still use the MLM analogue to repeated measures to yield an analysis similar in form and hypotheses evaluated to a traditional repeated measures analysis of variance. In short, the MLM analysis will still be preferred to standard application of the general linear model even if the measures must be used in a standard score metric. However, for measures that can be placed on a metric more appropriate for measuring change, it is advisable to do so.

This chapter was intended to provide a brief, nontechnical discussion of growth curve modeling along with an actual application of the method to an existing data set. The models examined were kept simple so that we could focus on interpretation of parameters and decisions about the modeling process. More extensive growth curve analysis of data from the CLS can be found in Francis et al. (1996). The

methods described in this chapter have undergone considerable development and extension in recent years to include models for growth in more than one domain (Willett & Sayer, 1994), models with nonnormally distributed outcomes (McCullagh & Nelder, 1989), and models that allow for simultaneous analysis of latent growth processes coupled with an underlying latent class model to represent finite mixtures of unknown populations (Khoo & Francis, 1998). It is hoped that the current chapter will assist researchers in their application of growth models in their own research and provide a background for extending their analysis toolbox to include the more complex models and analyses that were only briefly mentioned here.

ACKNOWLEDGMENTS. This work was supported in part by NICHD-sponsored grants HD21888 "Psycholinguistic and Biologic Mechanisms of Dyslexia" to Dr. Bennett Shaywitz; HD30995 "Early Interventions for Children with Reading Problems" to Dr. Barbara Foorman; and HD28172 "Detecting Reading Problems by Modeling Individual Growth" to Dr. David Francis. The authors wish to thank Drs. Sally and Bennett Shaywitz for generously granting us access to the CLS data for use in this chapter.

REFERENCES

- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101*, 147–158.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Efron, B. F., & Morris, C. N. (1979). Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association, 74*, 311–319.
- Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37–55.
- Francis, D. J., Fletcher, J. M., Stuebing, K. K., Davidson, K. C., & Thompson, N. R. (1991). Analysis of change: Modeling individual growth. *Journal of Consulting and Clinical Psychology, 59*, 27–37.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Fletcher, J. M., & Shaywitz, B. A. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*, 1–11.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Oxford University Press.
- Khoo, S., & Francis, D. J. (1998). Latent growth modeling of developmentally evolving constructs: Modeling sequential processes. Paper presented at the annual meeting of the American Education Research Association, April, San Diego.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). London: Chapman and Hall.
- Morris, C. N. (1983). Parametric empirical Bayes inference: Theory and applications. *Journal of the American Statistical Association, 78*, 47–65.
- Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 90*, 726–748.
- Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement, 20*, 335–343.
- Rogosa, D. R., & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychometrika, 50*, 203–228.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that

- dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, 326, 145–150.
- Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of Research in Education*, 15, 345–422.
- Willett, J. B., & Sayer, A. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116, 363–381.
- Zeger, S. L., Liang, K. Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics*, 44, 1049–1060.

II

Designing Research and Applying Measures and Methods

The second section of the volume considers critical issues in designing research in pediatric and clinical child psychology. Among the many issues that could be considered, one of the most common is determining which sampling procedures can significantly limit the generalizability of research findings. Consequently, investigators who conduct research with children and families with a wide range of populations need to understand the methodological implications of their sampling procedures and consider them in designing their studies and in analyzing and interpreting their findings. Careful consideration of potential sampling issues prior to conducting their studies should enhance investigators' abilities to make more informed decisions concerning the selection and recruitment of children in their studies. To address these issues, in Chapter 4, Drotar and Riekert describe concepts and strategies for sampling in research in pediatric and clinical psychology, sources of variation in samples, sample selection problems and strategies including recommendations for future research to help investigators reduce the impact of sampling problems on the integrity of their data.

As is true of other areas of psychology, measurement presents critical issues and problems for researchers who gather data from pediatric and clinical populations. Researchers in these fields are often in the position of developing new measures to address specific research questions as well as applying standardized measures to populations for which they have not been standardized. Consequently, investigators need to be conversant with a range of measurement issues to help them develop informed measurement strategies in their research programs. In Chapter 5, Overholser, Spirito, and DiFilippo present information concerning a wide range of measurement issues and problems that are commonly encountered in research in pediatric and clinical child psychology. Using liberal examples from their own research with pediatric and clinical populations, these authors describe strategies for measurement development in applied settings. These authors consider several methodological strategies that are relevant to such common research problems as applying established measures to new populations, modifying existing measures,

and developing new ones for specific research populations and issues in integrating data from multiple informants. Specific recommendations for training research assistants and managing research data are also considered by these authors.

One of the important needs in the fields of clinical child and pediatric psychology is to develop innovative methods of measurement that facilitate the development of new scientific knowledge. In Chapter 6, Quittner discusses the limitations of current approaches to assessment in the fields of pediatric and clinical child psychology, especially global assessment approaches and problematic measurement designs. Based on her own programmatic research with children with a range of chronic health conditions, Quittner presents an alternative model of assessment that involves a contextual approach to assess stressors and coping strategies. She takes readers through applications of the behavior analytic model, including situational analysis, item development, response enumeration, and response evaluation. One of the most useful features of this chapter is the detailed illustrations of measures that illustrate the application of the contextual approach to assessment including the Role-Play Inventory of Situations and Coping Strategies and the Daily Phone Diary Procedure. Quittner also gives valuable suggestions concerning training students in measurement issues and future research directions in measurement.

The next two chapters in this section concern measurement issues that are important to the development of the fields of clinical child and pediatric psychology but have been relatively neglected. These include application of qualitative methods and the consideration of cultural and ethnic influences. In Chapter 7, Krahn and Eisert present an interesting historical overview and rationale for the use of qualitative methods in the fields of pediatric and clinical child psychology. They consider relevant issues in using qualitative methods such as data collection methods, sampling analyses, and generalizability. Various alternatives for qualitative methods, for example, ethnography and ecological methods, are described along with criteria for using them in specific studies. One instructive feature of this chapter is the illustration of the application of qualitative methods in clinical child and pediatric psychology in several studies and populations, including the experiences and perceptions of parents who are raising a child with a conduct disorder, experiences of children whose lives had been disrupted by adoptions, social support in diabetes, and the author's work in qualitative assessment of parent satisfaction with clinical services. Finally, useful suggestions are given to expand research applications of qualitative research methods and training in the fields of child clinical and pediatric psychology.

In the final chapter in Section II (Chapter 8), Walders and Drotar describe the challenges that need to be met in order to enhance the inclusion of ethnic minority populations in research in pediatric and child clinical psychology, such as conceptual issues in the definition of race, design issues, such as the choice of between group and within group designs, and so forth. Concrete recommendations for sampling and accessing diverse populations and recruiting subjects from ethnic and minority populations are provided. Using examples from specific research programs in the United States and in other countries, this chapter also considers special issues in measurement that arise in conducting research with children from diverse cultures and their families. Finally, Walders and Drotar consider recommendations to enhance research training concerning cultural and ethnic issues for pediatric and clinical child psychologists.

4

Understanding and Managing Sampling Issues in Research with Children

DENNIS DROTAR and KRISTIN A. RIEKERT

Pediatric and clinical child psychologists often conduct research with heterogeneous clinical populations (e.g., children and adolescents with behavior disorders or chronic health conditions) that raise difficult sampling problems. Children and adolescents may be recruited from highly specialized populations and settings, for example, hospitals, that can result in biased samples (Berkson, 1946). Recruiting clinical populations can result in sample self-selection (Betan, Roberts, & McClusky-Fawcett, 1995) and attrition (Aylward, Hatcher, Stripp, Gustafson, & Leavitt, 1985), which may seriously limit generalizability of findings. Finally, in order to reduce heterogeneity of sample characteristics and carefully target their samples, investigators often restrict eligibility criteria, which also limits generalizability.

Given the above problems, investigators who conduct research with children in applied settings need to understand the methodological implications of sampling procedures and consider these issues in the design, presentation, and interpretation of their findings. Careful consideration of potential sampling problems will enhance investigators' abilities to make more informed decisions concerning selection and recruitment of children in clinical settings. Moreover, when it is not possible to prevent sampling bias, as is often the case, recognition of relevant sampling issues will facilitate more careful, complete description of sample selection procedures as well as more informed interpretation of findings (Betan et al., 1995; Drotar, 1994).

DENNIS DROTAR • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038. **KRISTIN A. RIEKERT** • Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123.

The purpose of this chapter is to help investigators understand the influence of sampling issues in research with pediatric and clinical child populations and develop a more informed approach to decision making concerning sample selection and description. The issues that will be considered in this chapter include concepts and strategies for sampling; sources of variation in samples; description of sample selection problems that are created by such factors as criteria, recruitment, self-selection, and retention; and strategies to reduce the impact of these problems on the integrity of data.

CONCEPTS OF SAMPLING

Definitions

To frame the present discussion of sampling issues, readers need to understand a few basic concepts and definitions concerning sampling. Interested readers are referred to texts on epidemiological methods for more extensive discussion of these and related concepts (Henry, 1990; Kleinbrum, Kupper, & Morgenstern, 1982; Verholst & Koot, 1992). In considering sampling issues, it is very helpful to distinguish among a hierarchy of populations about which conclusions can be drawn concerning a study question (Kleinbrum et al., 1982). The largest of these populations is the *external population*, which is defined as the group of individuals to which the study has not been restricted but to which one wishes to generalize the study findings. For example, if one is studying children with a chronic health condition such as cystic fibrosis (CF) at one hospital site, a relevant external population would be *all* children with CF at various sites throughout the country.

A second useful definition is the *target population*, which is defined as the collection of individuals of primary research interest from which one samples and to which one wishes to make statistical inferences concerning the study question. Using the above example, the target population would be all children with CF who meet criteria (e.g., age, duration of illness) at a particular site or multiple sites, depending on where the sample is recruited from. Finally, the *study population or sample* is defined as those children from whom data has been obtained. In the foregoing example, the study population would be those children with CF and their families who agreed to participate in the research and completed the study protocol.

Types of Error

Another set of important concepts in sampling concerns *error*, which is defined as the difference between the true population value of a measure for a target population and the estimate that is based on the data from the sample that is collected. Researchers also should differentiate between systematic error and random error. Kleinbrum et al. (1982) define *systematic error* as the difference between the true effect or parameter that the investigator is trying to estimate versus what is actually being estimated in the study. Systematic error or sampling bias (Henry, 1990) can result from various factors: For example, in the study described above,

systematic error could occur if children and adolescents with CF who had the most positive or adaptive patterns of psychological adjustment were recruited. The above scenario could occur in several ways, for example, if physicians selected only those families of children with CF who they thought would respond positively to the study and/or if the parents who chose to participate were those whose children had the most positive psychological adjustment, and so on. To the extent that such selective identification and participation occurs, the data obtained will be less representative of the target population and conclusions will be less valid for the target population.

In contrast to systematic error and sampling bias, *random error* is defined as a difference between the estimate computed from the study data and the parameter that is actually being estimated. In contrast to sampling bias, random error is not preventable and reflects sampling variation, the extent of which may depend on sample size and characteristics, for example, the variance in the dependent measure of interest. Random error, which is inherent in any study, affects the precision of the inferences concerning data but does not invalidate conclusions that can be drawn about the target population (Henry, 1990).

Type of Sampling Procedures

Probability Sampling

Depending on their scientific question and resources, investigators have the option of using various methods to select their samples (Fife-Schaw, 1995). *Probability samples* are selected so that every member of the target population has an equal chance or possibility of being included. Alternative types of probability sampling include *simple random*, where each member of the sample has an equal probability of being selected, or *stratified random sampling*, in which the sample is divided into subgroups (e.g., high vs. low socioeconomic status) to insure that subjects from each of the subgroups are equally included. In *cluster sampling*, the investigator selects at random a smaller number of clustering units, for example, schools, hospitals, and draws the sample from them. *Quota sampling* specifies the percentage of particular types of people, for example, gender, that needs to be included in the population. Finally, *theoretical sampling* selects groups who are thought to provide special insights or information concerning the investigator's questions. This method of sampling is most useful in hypothesis generation research. Each of the above methods of probability sampling has practical as well as methodological advantages and disadvantages (see Fife-Schaw, 1995, for more extensive discussion).

Population-based probability samples have special advantages for many research questions with clinical populations, for example, allowing for more representative results, eliminating geographic biases that could result from localized small-area samples, and giving more precise estimates of prevalence of conditions of interest, such as children with chronic illness (Newacheck & Taylor, 1992). Examples of population-based research in clinical child and pediatric psychology include the National Health Interview Survey on Child Health in the United States (Adams & Hardy, 1989), which has yielded data concerning the psychological impact of chronic health conditions (Silver, Westbrook, & Stein, 1998), and the

Ontario Child Health Study in Canada, which has yielded information on a wide range of populations, including children with mental and physical disorders (Cadman, Boyle, Szatmari, & Offord, 1987). The main disadvantages of population-based probability sampling strategies are the expense and the impracticality of using extensive outcome measures with large samples across an extensive population base.

Nonprobability Sampling

The second major type of sampling is a *nonprobability sample*, which is selected based on specific criteria and convenience (Henry, 1990). Many of the samples that are used by pediatric and clinical child psychologist researchers are nonprobability, convenience samples that are recruited based on availability and feasibility, that is, they are less expensive and more feasible to recruit than probability samples. The prevalence of nonprobability sampling in pediatric and clinical child psychology research requires investigators in these fields to carefully consider the potential impact on validity and generalizability of their findings. Probability sampling can heighten the impact of confounding factors such as socio-economic status that could account for variance in dependent measures. For example, investigators who recruit children with asthma from a hospital that serves economically disadvantaged children and their families might be expected to obtain less optimal psychological adjustment in this sample than in a sample that is recruited from a suburban practice. Hospital-based samples of children with asthma also might be expected to demonstrate more behavioral problems than normative samples on which widely used measures such as the Child Behavior Check List (CBCL) are standardized (Achenbach, 1991). Consequently, findings from such uncontrolled studies can be erroneously attributed to the specific influence of the child's asthma rather than demographic characteristics and risk factors associated with a specific sample (Perrin, Stein, & Drotar, 1991).

SAMPLE SELECTION, RECRUITMENT, AND ATTRITION BIAS

One of the most important problems in conducting research with children and adolescents, especially in applied settings, is bias concerning sample self-selection. Selection bias is a potential problem in any study but is of particular concern in studies that recruit relatively small, nonprobability convenience samples of pediatric and clinical child populations from single sites.

The Impact of Selective Participation and Attrition in Research with Children and Families

Sample biases occur for several reasons. One common source is self-selection based on parental and/or child consent. Studies of various populations have consistently documented that children of parents who refused permission for their participation in research have greater psychosocial difficulties, including more problematic relationships with peers (Beck, Collins, Overholser & Terry, 1984; Frame & Strauss, 1987), academic problems (Frame & Strauss, 1987; Noll, Zeller,

Vannatta, Bukowski, & Davies, 1997), attentional difficulties and depression (LaGreca & Silverman, 1993), and aggressive behavior (Noll et al., 1997) than children whose parents gave them permission to participate.

Similar sample participation biases also might occur for pediatric chronic illness samples, although the data are not extensive. Riekert and Drotar (1998) found that the families of adolescents with a chronic illness such as diabetes who participated in research on treatment adherence demonstrated higher levels of adherence to medical treatment than those who did not participate.

Significant selection biases also may operate with respect to children's agreement to participate in research. Weinberger, Tublin, Ford, and Feldman (1990) found that boys from 33 classrooms who were nominated by their peers as high in distress and moderate or low in self-restraint were significantly less likely than other boys to volunteer to take part in an in-class survey.

Participation bias also can be a significant problem in intervention studies with children and families. For example, in a clinical trial designed to assess the efficacy of a comprehensive early intervention program in improving the health and development of low-birth weight infants, Constantine, Haynes, Spiker, Tackett, and Constantine (1993) found that mothers of lighter-weight infants were more likely to enroll than were mothers of heavier-weight infants. Moreover, African Americans and Hispanics were more likely to participate than were whites.

Stein, Bauman and Ireys (1991) found that parents of infants in a neonatal intensive care unit who participated in an intervention were more likely to have infants with lower gestational age and with longer hospital stays than members who did not participate. Moreover, participants tended to have higher incomes than nonparticipants. These authors concluded that many of the mothers whom the investigators viewed as most likely to be helped by the intervention were least likely to participate.

Participation bias is not the only kind of sampling problem for researchers in pediatric and clinical child psychology. Selective attrition also can pose significant problems in descriptive studies of pediatric and clinical child populations, for example, low-birth weight infants, (Aylward et al., 1985) as well as treatment studies (Frankel & Simmons, 1992).

Describing the Impact of Selective Participation

Despite the considerable evidence for sample bias that is attributable to selective participation, pediatric and clinical child psychologist researchers generally have not documented the impact of these problems nor have they provided comprehensive descriptions of their samples. Betan et al.'s (1995) content analysis of published articles in the *Journal of Clinical Child Psychology* and the *Journal of Pediatric Psychology* found that the mean reported consent rate defined as the proportion of those families who were initially approached who were represented by the final sample was only 65.9%. While such consent rates are not necessarily synonymous with sample bias, they raise questions about representativeness of the samples that are collected. Unfortunately, only a minority of published articles (42%) in the above journals reported consent rates let alone information about the characteristics of the families who give consent versus those who did not (Betan et al., 1995).

Considering the Impact of Selective Participation on the Validity of Research Findings

While many researchers recognize that their research results cannot necessarily be generalized to nonparticipants, the potential impact of selective participation on the internal validity of study findings is less commonly considered. For example, the evaluation of the efficacy of preventive intervention with a high-risk sample may be compromised by selective enrollment of subjects based on their risk characteristics (Stein et al., 1991). Consider a study of an intervention designed to prevent child maltreatment. If a disproportionate number of families who had relatively low risk for maltreatment participated, a larger sample than initially anticipated would be necessary to demonstrate an intervention effect (Stein et al., 1991).

Selective participation also may affect the quality of data concerning theories of intervention process. By definition, selective participation means that an intervention program has reached some but not all of the intended population. Consequently, an intervention program that is found to be ineffective when applied to a biased sample may be erroneously interpreted as discrediting the theoretical model on which the intervention was based. However, the model eventually may prove to be correct if it is applied to the intended target population (Stein et al., 1991).

UNDERSTANDING SOURCES OF SAMPLING VARIATION IN CLINICAL POPULATIONS

Investigators who conduct research with heterogenous clinical populations of children need to understand the sources of variation of their samples, which can reflect such factors as individual differences in etiology, duration, and severity of clinical conditions, as well as setting characteristics, practitioner referral patterns, and medical and/or psychological treatments that are received by research participants.

Condition-Related Variation

The extraordinary variation in psychological outcomes that is attributable to individual differences in various clinical conditions can have a powerful influence on the nature and interpretation of findings based on clinical research. For this reason, investigators should familiarize themselves with the specific sources of variation in their samples. Interested readers are referred to discussions on the impact of condition-related variation for various clinical problems (Drotar, 1990; Karney et. al., 1995; Kopp & Krakow, 1982; Lamphear, 1986; Widom, 1988). In the next section we consider the potential impact of etiology, duration, and severity of clinical conditions.

Etiology of Conditions

The heterogeneous etiologies of many clinical problems may have a profound influence on children's psychological outcomes. This is a special problem in

research involving children with nonspecific clinical problems such as failure to thrive (FTT) (Drotar, 1990) and child abuse (Lamphear, 1986), which have widely divergent etiologies. Unless they are mindful of such variation, investigators may draw misleading conclusions about the sample as a whole or fail to identify important subtypes. For example, FTT can result solely from inadequate parental knowledge of children's nutritional needs or reflect multifaceted parent-child relationship and family organizational problems (Drotar, 1990). Consequently, investigators who study the characteristics of children who fail to thrive may reach very different conclusions, depending on which particular subgroups of children with this condition are recruited.

Duration of Conditions

The duration of a condition may be especially important in evaluating the outcomes of children with pediatric chronic illness as well as children with mental disorders. Data concerning child and family psychological distress may be heavily influenced by the timing of measurement in relation to the course of chronic physical illness (Kovacs, Brent, Steinberg, Paulauskas, & Reid, 1986). For example, shortly after the diagnosis of a chronic health condition, children tend to report relatively high levels of emotional distress that diminish thereafter (Kovacs et al., 1986). On the other hand, the incidence of problems such as posttraumatic stress symptoms is higher in long-term survivors of conditions such as childhood cancer (Kazak et al., 1997).

The above findings suggest that data concerning the stressful impact of chronic health and mental health conditions in children might be very difficult to ascertain in samples that include children with conditions that vary widely in duration, as is typically the case. For example, samples that contain disproportionately large numbers of chronically ill children with a recent onset of their illness might be biased to detect increased levels of psychological distress. Moreover, divergent findings in studies that are conducted in different settings might be explained by site-specific differences in onset, duration, and illness course in local samples of children (Drotar, 1994).

Condition Severity

Another important source of variation in children's psychological status is the severity of a child's clinical problem. Depending on the research question, the presence of children who are diagnosed with the identical mental or physical problem but who have very different frequencies of symptoms could obscure findings and/or lead to misleading inferences. For example, children diagnosed as having attention deficit hyperactivity disorder (ADHD) demonstrate extraordinary variation in the frequency and severity of attentional problems as well as the presence or absence of comorbid problems (Biederman, Farone, & Lapey, 1992). Consequently, the impact of ADHD on the family would be much greater for a sample that included a majority of children with severe ADHD-related symptoms and/or comorbid problems versus a sample who had a broad range of severity of ADHD symptoms.

Impact of Setting-Related Variation

Differences in settings from which children are recruited can have a powerful influence on children's psychological outcomes. Relevant setting characteristics include referral patterns, practitioners' selection of patients to participate in studies, the nature of medical and psychological treatment provided in individual settings, and typical family demographic characteristics at different sites.

Referral Patterns for Clinical Problems

Children with comparable medical conditions and/or psychological disorders may have very different rates or types of psychological problems as a function of the referral patterns in different settings (Berkson, 1946; Friman, 1993). For example, children with conditions such as FTT who are primarily recruited from ambulatory care settings would be expected to demonstrate very different (e.g., less severe and/or complex psychological and physical growth deficits) than those who are recruited from children who are hospitalized for the same condition (Drotar, 1990).

Investigators also need to consider how characteristics may influence who participates in their studies. In conducting research in school settings, researchers may encounter some school administrators who do not allow research to be conducted in their schools. Selected characteristics of schools associated with participation or nonparticipation may be reflected in characteristics of study samples: For example, administrators of science and math magnet schools may be more willing to participate in research than those of nonmagnet schools because their curriculum emphasizes research principles and provides students with opportunities to participate in research. On the other hand, the children who attend a science and math magnet school (and their families who chose to send them there) may have very different psychological characteristics (e.g., motivation for learning) than the larger population of students from which they are drawn.

Sample Bias from Practitioners' Selection of Patients for Research

Another source of sample bias stems from practitioners who select patients on the basis of subjective criteria, that is, their perceived likelihood of cooperating with the study procedures (see Chapter 12, this volume). Suppose that a psychological researcher who wants to recruit a sample of children with cancer asks his hematologist colleagues to contact their patients who meet a certain criteria, for example, duration of the condition. However, in an effort to help the psychologist, the well-meaning hematologists decide to contact only the patients who are judged as coping well because they believe that they will give the "best data." The colleagues "helpfulness" could result in selection of a sample that is biased in the direction of better-adjusted families. Moreover, unless the researcher knew about his colleagues' strategies for patient selection, the bias would not be detected.

Type of Medical and Psychological Treatments Provided to Patients

Children who are recruited from clinical settings also receive a range of medical and psychological treatments that may affect their psychological out-

comes. While the presence of medical treatment and psychosocial support services is rarely acknowledged as a source of variation in psychological outcomes of clinical populations, it may have important effects on research findings in both intervention and descriptive studies (Drotar, 1989). Consider a study of the psychological outcomes and coping abilities of children with sickle cell anemia: What if most children who participated in the study also received a comprehensive pain management program, but for a variety of reasons (e.g., transportation problems and poor compliance) nonparticipants were less likely to receive this intervention. To the extent that the comprehensive pain management program improved children's adjustment and coping, this could influence the number of children who were found to show positive coping and limit the generalizability of findings to other samples who did not receive such intervention.

Family-Demographic Characteristics of Samples Recruited in Different Settings

Family environmental and demographic characteristics (e.g., family economic and relationship resources, parental education and occupation, race, culture, or ethnicity) associated with patient populations at different settings also can vary considerably and have a significant influence on psychological outcomes of clinical populations. (Drotar, 1994). One interesting example was reported from the multisite Infant Health and Development Program (IHDP). In this study, the mean Stanford-Binet IQs of preschool children with histories of low birth weight who received comprehensive early intervention varied substantially across sites, from 84.7 to 102.5 (Infant Health and Development Program, 1990). Studies at single sites might have reached very different and perhaps equally erroneous conclusions concerning the relationship of birth weight to children's cognitive development.

THE IMPACT OF INVESTIGATORS' DECISIONS CONCERNING SAMPLE SELECTION AND RECRUITMENT

Inclusionary Criteria

In order to reduce the heterogeneity of their samples and facilitate more accurate tests of their study questions, investigators need to carefully set criteria for including subjects in their samples. Such critical decisions exert considerable influence on sample characteristics, and hence on the generalizability of findings. For example, one common exclusion in psychological studies of childhood chronic illness is children with multiple chronic conditions. Such exclusion is warranted for tests of some research questions, especially if the presence of a comorbid chronic condition obscures data obtained from children with the chronic condition of primary interest. However, because children with more than one chronic illness are excluded from most samples, little is currently known about their psychological functioning. While a relatively small number of children have two or more chronic conditions, these children may have more mental and physical health problems and use more health services than other children (Newacheck & Stoddard, 1994).

Incidence and Prevalence as Selection Criteria

Investigators' decisions to select subjects on the basis of incidence (new cases) or prevalence (existing cases) raise sampling problems that are not often considered in psychological research. For example, selecting cases by incidence (e.g., children who are newly diagnosed with depression) allows the investigator to gather data on subjects who are starting from the same time point in the course of their condition. This strategy is useful for research on factors that influence children's response to diagnosis of a chronic condition. On the other hand, this selection strategy may not be feasible or necessary for other studies.

In most research in pediatric and clinical child psychology, selection of cases is more commonly done on the basis of prevalence, which poses problems for interpretation of data that investigators should consider. For example, the duration of many childhood illnesses (e.g., diabetes, asthma) is often highly but not perfectly correlated with the child's age. Consequently, investigators who select heterogeneous age groups of children on the basis of their prevalence will not be able to determine whether psychological outcome data obtained from such samples reflect the influence of age or illness duration.

The impact of a natural history of a condition, including prognosis and mortality rates or psychological outcomes also should be considered. For example, assessing depression in a study of young adult survivors of CF may underestimate the prevalence of depression, given the following scenario; for example, if depression contributed significantly to poorer adherence to treatment, which, in turn, enhanced the rate of physical deterioration and mortality.

RECOMMENDATIONS

In this section, we summarize recommendations to help researchers develop strategies to limit the impact of sampling problems on their data, to present their sampling methods effectively, and to evaluate the impact of sampling issues on their findings.

Anticipate Sampling Problems in Designing Studies

In designing and implementing their research, investigators should consider strategies that limit the impact of sampling problems and enhance generalizability of findings. As shown in Table 1, these strategies include making informed decisions concerning inclusionary and exclusionary criteria, establishing control of sampling and recruitment procedures, reducing the impact of sampling influences on dependent measures by using appropriate controls, limiting comparisons with test norms, recruiting underrepresented populations, and preventing attrition.

Make Informed Decisions about Inclusionary and Exclusionary Criteria

Researchers should carefully consider the costs versus benefits of inclusionary and exclusionary criteria that are appropriate for their research questions and populations. In making their decisions about study criteria, investigators should consider such factors as the heterogeneity of the condition of interest, presence or

Table 1. Strategies to Enhance Generalizability of Psychological Research with Child Populations

-
1. Make informed decisions concerning inclusionary and exclusionary criteria
 - a. Consider costs and benefits of criteria that are used
 2. Establish control over sampling and recruitment strategy
 - a. Standardize procedures
 - b. Obtain access to entire sample
 3. Present a clear and comprehensive description of sampling methods and sampling characteristics
 - a. Describe criteria
 - b. Describe methods by which sampling was obtained
 4. Use methods to quantify impact of sampling bias
 5. Use data-analytic procedures to limit the impact of sampling influences
 - a. Document relationship of subject characteristics to outcomes
 - b. Adjust for cofactors
 - c. Utilize meta-analyses
 6. Design studies to limit the impact of sampling influences
 - a. Choose an approximate control group
 - b. Limit comparisons with test norms
 - c. Recruit underrepresented populations
 - d. Replication
 - e. Multisite studies
 - f. Population-based designs
 - g. Theory-based research
 - h. Conduct studies of subgroups
-

absence of associated risk factors, use of incidence and prevalence as criteria, and other associated factors such as age, and so forth, and available sample size. Depending on their research questions, investigators may prefer to use restrictive criteria in order to reduce variance in their outcomes and/or choose to identify a relevant subgroup as opposed to sampling strategy that involves a broad cross-section of the population of interest.

Because restrictive inclusionary criteria will severely limit the generalizability of their findings, investigators should base their decisions on a thorough understanding of the implications of alternative criteria in reference to their research questions. For example, investigators who are interested in the process of psychological adaptation to a newly diagnosed chronic illness would obviously restrict their sample to children and adolescents with a recent diagnosis. On the other hand, recently diagnosed children logically might be excluded in research concerning adherence to medical treatment regimens in order to provide a more informed test of the factors that affect longer-term adherence to a continuous treatment regimen. Because children with a recent diagnosis of a child's illness and their families are learning an unfamiliar and complex treatment regimen, their management of treatment adherence may be very different than children and families who have had to cope with the demands of their conditions for much longer periods of time.

Establish Control of Sampling and Recruitment Procedures

Wherever it is feasible to do so, researchers who collect data concerning children and families in clinical settings should assume direct control over the

sampling strategy and procedures that are used to select and recruit their samples, rather than rely solely on practitioners to determine who gets into their study versus who does not (see Chapter 12, this volume). For example, in those settings where data are available concerning relevant characteristics of the sample (e.g., age, family financial status, date of diagnosis, etc.), investigators can independently ascertain whether a potential subject meets criteria by reviewing all the patient records of potentially eligible children.

Investigators who conduct research in clinical settings also should anticipate that practitioners' referrals may be biased and impress on their practitioner colleagues the necessity and importance of recruiting as unbiased a sample as possible. In order to ascertain and prevent referral bias, researchers also should work closely with their practitioner colleagues to understand how they identified patients for the study.

Choose an Appropriate Control-Comparison Group

An ideal control group consists of individuals who are selected from a population that is similar to the study group in their distribution of demographic characteristics, and so forth, other than the independent variable under consideration. However, clinical populations tend to be quite heterogeneous. Study subjects who are recruited from clinical settings may come from the surrounding neighborhood, live further from the setting in the same city, or reside in other towns or cities or even out of state. Consequently, it may be very difficult for investigators to determine the reference population from which such heterogeneous study samples are drawn, and hence to select an appropriate control group.

One solution to the dilemma would be to select control subjects who reside in the same neighborhood as the study subjects. One method of doing this is to ask the study subject to identify a "best friend" who would serve as a control (Noll et al., 1996). It is assumed that a friend would be similar to the study subject on demographic and social variables and would be more willing to participate in the study than a stranger. This solution, however, may or may not be feasible, depending on resources for data collection. Other options are to recruit controls from pediatric practices or schools.

Limit Comparisons with Test Norms

Depending on the sample and research question, comparisons of a sample's responses to normative data from tests such as the CBCL (Achenbach, 1991) may be misleading (Drotar, Perrin, & Stein, 1995; Perrin et al., 1991). Other differences between study versus normative samples may reflect the time period when the normative data were collected (i.e., cohort effects) and the geographic distribution of the study sample compared to the normative sample. To avoid these problems, investigators may want to limit their use of comparisons with test norms.

Sample and Recruit Underrepresented Populations

Another general recommendation for investigators is to sample and recruit underrepresented populations. This is especially important in conducting research

with populations that have been underrepresented in research in pediatric and clinical child psychology, for example, ethnic and minorities children and adolescents (see Chapter 8, this volume).

Limit Selection Bias and Attrition via Strategies of Recruitment and Retention

In dealing with selective bias and attrition, prevention is the best strategy. Consequently, investigators should use a range of methods to enhance consent rates and reduce dropout rates, such as incentives for participants, persistent contacts by project staff, and so on, especially when working with participants who are difficult to recruit and maintain in research. For more detail, see Chapter 12, this volume.

Present a Clear and Comprehensive Description of Sampling Methods and Sample Characteristics

While relevant sources of sampling variation often cannot be completely controlled, they can be described thoroughly and carefully considered in interpreting findings. We concur with others' recommendations (Betan, et al., 1995; Karney et al., 1995) that investigators should describe their sampling procedures and characteristics in more detail than has been customary. We present guidelines for such description in Table 2 and Fig. 1. See Begg et al. (1996) for a detailed description of sampling and participation in randomized controlled intervention trials.

A detailed description of selection criteria (including rationale for inclusions and exclusions, recruitment procedures, and relevant characteristics of the settings from which participants were recruited) will help other researchers to understand the procedures that were utilized, and hence evaluate the findings. Moreover,

**Table 2. Description of Sampling and Subjects:
What Should Be Included in Published Reports**

-
1. Describe and provide rationale for inclusions/exclusionary criteria that are implemented
 - a. Describe each specific criterion in detail
 - b. Provide a rationale for criteria, e.g., why were they chosen?
 2. Describe setting from where sample was drawn
 - a. Type of setting, e.g., hospital, child guidance clinics
 - b. Characteristics of medical and psychosocial treatments given to the sample
 3. Describe subjects who fit criteria and who were approached for consent, including:
 - a. Relevant characteristics of refusals if known
 - b. Age, gender, etc., and similarities and differences between refusals and participants
 4. Describe characteristics of attrition sample in prospective studies
 - a. Demographic or characteristics
 - b. Initial measures on attrition sample
 - c. Analysis of similarities and differences between attrition subjects and sample
 5. Describe sample
 - a. Number
 - b. Relevant demographic characteristics, e.g., age, sex, social class
 - c. Relevant clinical characteristics, e.g., severity of condition, chronicity, type of treatment received by subjects
-

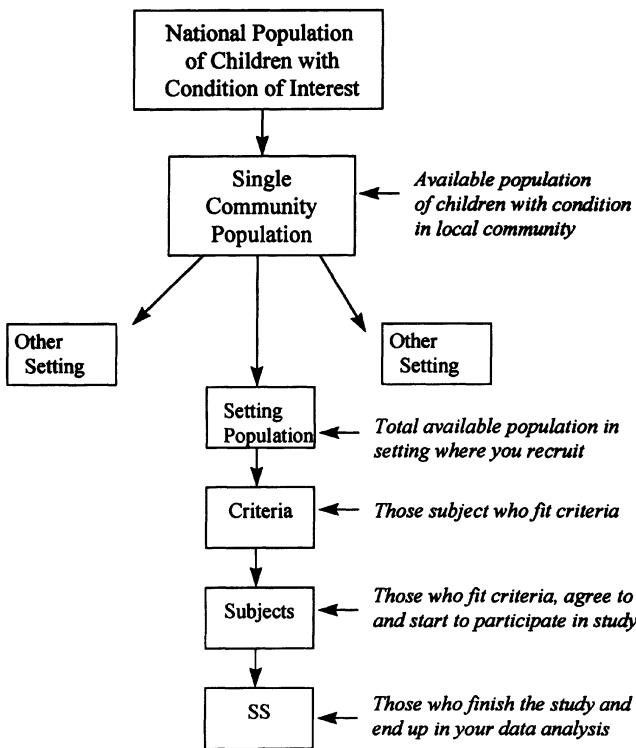


Figure 1. Sampling considerations in pediatric psychology.

children who fit study criteria and were approached for consent but did not participate should be described and compared to participants with respect to demographic characteristics and other available information (Riekert & Drotar, 1999). Similarly, the characteristics of subjects who began the study but did not complete data collection should be described and compared with that of participants. Such comprehensive descriptions of study participants would facilitate replication of the study by others, enable other researchers to consider and evaluate study findings more effectively, as well as to compare data gathered at different sites (Betan et al. 1995). Such detailed descriptions also facilitate researchers' interpretation of their data.

Use Data Analytic Procedures to Clarify and Limit the Impact of Sampling Influences

A number of data analytic strategies are available to investigators to help them clarify and control for the impact of sampling influences. Several of these are described in the next section.

Statistical Procedures to Adjust for Relevant Cofactors

Investigators should evaluate how clinically relevant family and other characteristics may have influenced their primary dependent measures. To adequately address some research questions (e.g., determining the effects of high lead levels on developmental outcomes), formal statistical procedures (e.g., multiple regression) to adjust for influential cofactors may be critical. Greene and Ernhart (1991) have described a useful decision-making strategy to determine whether to control statistically for cofactors and present methods for conducting such analyses.

Meta-analyses

Statistical methods such as meta-analysis also can be effectively used to summarize existing data sets, to assess generalizability of findings, and ascertain how patterns of study findings vary as a function of subject (e.g., type of condition) and/or setting characteristics (Bangert-Drowns, 1986). For example, Lavigne and Faier-Routman's (1991) meta-analysis of children with chronic health conditions found that the impact of conditions that involved the central nervous system was greater on children's psychological adjustment than those that did not.

Methods to Assess Sampling Bias due to Selective Participation and/or Attrition

Several data analytic options are available to investigators to manage selection bias and attrition. In some cases, it is possible to sometimes use statistical procedures to estimate the actual population statistics that are based on the biased sample. For example, in a study designed to estimate the prevalence of depression in a mental health center sample, Frank, Schulberg, Welch, Sherick, and Costello (1985) used a screening interview to obtain information about subjects' demographics, characteristics, and depression status. These researchers followed the screening interview with a more in-depth diagnostic interview and found that many people either declined participation or failed to keep an appointment for the longer interview. The investigators then implemented a two-step statistical procedure as suggested by Welch, Frank, and Costello (1983) to correct for sampling bias due to selective participation in a second interview. This procedure involved an a priori identification of variables that were expected to predict the pattern of self-selection.

Researchers also should understand that their methods of sampling also may influence their ability to estimate the impact of sampling bias on their data. For example, because they provide data on all potential members of a sample, whenever it is feasible to do so sampling so that comprehensive records of an available sample can be obtained is preferable to other sampling procedures, for example, via newspaper advertisements, for which sampling bias cannot be estimated (Karney et al., 1995).

Statistical Methods to Test for Biases due to Attrition

To determine whether attrition has biased a sample, Jurs and Glass (1971) suggested performing a series of multivariate analyses of variance (MANOVAs) using

attrition status and experimental treatment group as independent variables and the primary outcome variables as dependent measures. Other methods for assessing and managing attrition that are possible in some studies include endpoint analysis (Chassan, 1979), in which the last assessment for each subject who drops out of the study is his or her endpoint value; time-controlled analysis (Cook & Campbell, 1979) in which data can be collected on subjects after they have dropped out of treatment; covariance techniques to adjust for inequalities between groups of variables that influence attrition; and statistical methods to replace missing data. See Flick (1988) for a useful description of the costs versus benefits of these and other procedures.

STRATEGIES TO ENHANCE GENERALIZABILITY OF RESEARCH WITH CHILDREN

Researchers in pediatric and clinical child psychology should consider using a range of strategies to enhance the generalizability of data obtained from clinical populations in various settings.

Replicate Research Findings

Replicating findings obtained with samples and methods in one setting in a different setting is a classic method to enhance generalizability of findings (Campbell & Stanley, 1967) that has been relatively neglected in research in clinical child and pediatric psychology. We believe that replications warrant greater attention from investigators and high priority for publication from editors in these fields.

Use Multisite and Population-Based Studies

The overreliance on convenience sampling in research in pediatric and clinical child psychology indicates that multisite and population-based studies need to be used much more frequently than they have been (see Chapter 13, this volume). Successful multisite studies such as the Infant Health and Development Program (1990) with low-birth weight infants and the Growth and Development Study in hemophilia (Sinois et al., 1998) have generated valuable information concerning specific populations. Moreover, population-based studies have been used to great advantage to address research questions such as the incidence and prevalence of mental disorders among children with chronic health conditions and their families (Cadman et al., 1987).

Conduct Studies of Subgroups of Clinical Populations

Another useful strategy for future research is to identify subgroups of children with clinical problems that differ on psychologically relevant outcomes and establish the validity of these subgroups. For example, child and family noncompliance with medical treatment is an important but poorly defined problem in child health psychology. In particular, the heterogeneous patterns of problematic adherence require better definition. Using family members' descriptions of critical incidents, Koocher, McGrath, and Gudas (1990) identified patterns of nonadherence to treat-

ment regimens among children and adolescents with CF, including: (1) inadequate knowledge; (2) psychosocial resistance (e.g., control struggles with parents, cultural pressures, striving for normality); and (3) educated nonadherence (e.g., a rational decision to disregard certain elements of treatment). Research is needed to validate these subtypes, establish their clinical significance, and develop methods to assess and identify them.

Test Theories to Describe Processes That Generalize across Samples

One of the most important ways to enhance the generalizability of research in pediatric and clinical child psychology is to test theoretical models of critical psychological processes across a range of populations. While theory-based research has been relatively neglected, especially in pediatric psychology research (Wallander, 1992), examples of this approach include tests of conceptual models of risk and resistance factors concerning the psychological adaptation in chronic illness populations (Thompson, Gil, Burbach, Keith, & Kinney, 1993; Wallander et al., 1989).

Identify Factors That Affect Sampling Bias

Researchers who work with children and families should continue to examine the factors that contribute to sampling bias in various populations (Flick, 1988). Such data can help to illuminate potentially important issues, for example, patterns of personality of children who refuse to participate in research (Weinberger et al., 1990) or the high-risk developmental status of untestable subjects (Bathurst & Gottfried, 1987).

Future research also should identify potential variables that mediate and/or moderate the relationship between participation status and data concerning health and psychological outcomes. For example, Janus and Goldberg (1997) found that lower maternal age at the birth of a child, lower paternal education, and greater disorganization of mother–infant attachment were all related to greater sample attrition for samples including children with CF and coronary heart disease.

Develop Recommendations to Improve the Quality of Sampling in Research with Children and Families

To enhance generalizability of individual research efforts, those of us who are actively involved in research with children and families should develop consensus, state-of-the-art recommendations concerning sampling issues, criteria, and recruitment strategies that have been found to be necessary and feasible for various clinical populations. Spiker and colleagues' description of the costs and benefits of decisions made in designing the Infant Health and Development Program (Spiker, Kraemer, Scott, & Gross, 1992) is an excellent example of such work. Others should follow their lead.

REFERENCES

- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington: University of Vermont Press.

- Adams, P. F., & Hardy, A. M. (1998). Current estimation from the National Interview Survey: United States Health. *Vital and Health Statistics, Series 10*, No. 173. Hyattsville, MD: National Center for Health Statistics.
- Aylward, G. P., Hatcher, R. P., Stripp, R., Gustafson, N. F., & Leavitt, L. A. (1985). Who goes and who stays? Subject goes in multicenter longitudinal follow-up study. *Journal of Developmental Behavioral Pediatrics, 6*, 3-8.
- Bangert-Drowns, R. L. (1986). Review of psychological developments in meta-analytic methods. *Psychological Bulletin, 99*, 383-399.
- Bathurst, K., & Gottfried, A. W. (1987). Untestable subjects in child development research: Developmental implications. *Child Development, 58*, 1135-1144.
- Beck, S., Collins, L., Overholser, J., & Terry K. (1984). A comparison of children who receive and who do not receive permission to participate in research. *Journal of Abnormal Child Psychology, 12*, 573-580.
- Begg, C., Cho, M., Erstwood, S., Horten, R., Moher, D., Oikin, I., Pitkin, R., Rennie, D., Schulz, K. F., Simel, D., & Stroupe, D. R. (1996). Improving the quality of reporting in randomized controlled trials. The CONSORT statement. *Journal of the American Medical Association, 276*, 637-639.
- Berkson, J. (1946). Limitation of the application of the fourfold table analysis to hospital data. *Biometrics Bulletin, 2*, 47-53.
- Betan, E. J., Roberts, M. C., & McCluskey-Fawcett, K. (1995). Rates of participation for clinical child and pediatric psychology research issues in methodology. *Journal of Clinical Child Psychology, 24*, 227-236.
- Biederman, J., Farone, S. V., & Lapey, K. (1992). Comorbidity of diagnosis in attention-deficit hyperactivity disorder. *Child and Adolescent Psychiatric Clinics of North America, 1*, 335-360.
- Cadman, D., Boyle, M., Szatmari, P., & Offord, D. R. (1987). Chronic illness, disability, and mental and social well being: Findings of the Ontario Child Health Study. *Pediatrics, 79*, 805-813.
- Campbell, D. T., & Stanley, J. L. (1967). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Chassan, J. B. (1979). *Research design in clinical psychology and psychiatry* (2nd ed.). New York: Irvington Publishers.
- Constantine, W. L., Haynes, C. W., Spiker, D., Tackett, K. K., & Constantine, N. K. (1993). Recruitment and retention in a clinical trial for low birth weight, premature infants. *Journal of Developmental and Behavioral Pediatrics, 14*, 1-7.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally College Publishing.
- Drotar, D. (1989). Psychological research in pediatric settings: Lessons from the field. *Journal of Pediatric Psychology, 14*, 69-74.
- Drotar, D. (1990). Sampling considerations in research with nonorganic failure to thrive. *Journal of Pediatric Psychology, 15*, 255-272.
- Drotar, D. (1994). Psychological research with pediatric conditions: If we specialize can we generalize? *Journal of Pediatric Psychology, 19*, 403-414.
- Drotar, D., Perrin, E. C., & Stein, R. E. K. (1995). Methodological issues in using the Child Behavior Checklist and its related instruments in clinical child psychology research. *Journal of Clinical Child Psychology, 24*, 184-192.
- Fife-Schaw, C. (1995). Surveys and sampling issues. In G. M. Breakwell, S. Hammond, & C. Fife-Schaw (Eds.), *Research methods in psychology* (pp. 99-115). Thousand Oaks, CA: Sage.
- Flick, S. N. (1988). Managing attrition in clinical research. *Clinical Psychology Review, 8*, 499-515.
- Frame, C. L., & Strauss, C. C. (1987). Parental informed consent and sample bias in grade school children. *Journal of Social and Clinical Psychology, 5*, 227-236.
- Frank, R. G., Schulberg, H. C., Welch, W. P., Sherick, H., & Costello, A. J. (1985). Research selection bias and the prevalence of depressive disorders in psychiatric facilities. *Journal of Consulting and Clinical Psychology, 53*, 370-376.
- Frankel, F., & Simmons, J. Q. (1992). Parent behavioral training. Why and when some parents drop out. *Journal of Clinical Child Psychology, 21*, 322-330.
- Friman, P. C. (1993, April). Assessing the relationship between common child referral problems and psychopathology: The possible influence of Berkson's bias. Presented at the *Fourth Florida Conference on Child Health Psychology*, Gainesville, FL.
- Greene, T., & Ernhart, C. (1991). Adjustment for cofactors in pediatric research. *Journal of Developmental and Behavioral Pediatrics, 12*, 378-386.
- Henry, G. T. (1990). *Practical sampling*. Newbury Park, CA: Sage.

- Infant Health and Development Program. (1990). Enhancing the outcomes of low-birthweight, pre-mature infants: Results of a multisite, randomized trial. *Journal of American Medical Association*, 263, 3035–3042.
- Janus, M., & Goldberg, S. (1997). Factors influencing family participation in a longitudinal study: Comparison of pediatric and healthy samples. *Journal of Pediatric Psychology*, 22, 245–262.
- Jurs, S. G., & Glass, G. V. (1971). The effect of experimental mortality on the internal and external validity of the randomized comparative experiment. *Journal of Experimental Education*, 40, 62–66.
- Karney, B. R., Davila, J., Cohan, C. L., Sullivan, K. T., Johnson, M. D., & Bradbury, T. N. (1995). An empirical investigation of sampling strategies in marital research. *Journal of Marriage and the Family*, 57, 909–920.
- Kazak, D. E., Barakat, L. P., Meeske, K., Christakis, D., Meadows, A. T., Casey, R., Penati, B., & Stuber, M. L. (1997). Post-traumatic stress, family functioning, and social support in survival of childhood leukemia. *Journal of Consulting and Clinical Child Psychology*, 68, 620–629.
- Kleinbrum, D. G., Kupper, L. L., & Morgenstern, H. (1982). *Epidemiologic research principles and quantitative methods*. Belmont, CA: Lifetime Learning Publications.
- Koocher, G. P., McGrath, M. L., & Gudas, L. J. (1990). Typologies of nonadherence in cystic fibrosis. *Journal of Developmental and Behavioral Pediatrics*, 11, 353–358.
- Kopp, C., & Krakow, J. B. (1982). The issue of sample characteristics: Biologic at risk or developmentally delayed infants. *Journal of Pediatric Psychology*, 7, 361–375.
- Kovacs, M., Brent, D., Steinberg, T. F., Paulauskas, S., & Reid, J. (1986). Children's self-reports of psychological adjustment and coping strategies during the first year of insulin-dependent diabetes mellitus. *Diabetes Care*, 9, 472–479.
- LaGreca, A. M., & Silverman, W. C. (1993). Parent reports of child behavior problems. Bias in participation. *Journal of Abnormal Child Psychology*, 21, 89–101.
- Lamphier, V. S. (1986). The psychosocial adjustment of maltreated children. Methodological limitations and guide lines for future research. *Child Abuse and Neglect*, 10, 63–69.
- Lavigne, J. V., & Faier-Routman, J. (1992). Psychological adjustment to pediatric physical disorders: A meta analytic review. *Journal of Pediatric Psychology*, 17, 133–158.
- Newacheck, P. W., & Stoddard, J. J. (1994). Prevalence and impact of multiple childhood chronic illnesses. *Journal of Pediatrics*, 124, 40–48.
- Newacheck, P. W., & Taylor, W. R. (1992). Childhood chronic illness: Prevalence severity, and impact. *American Journal of Public Health*, 82, 364–370.
- Noll, R. B., Vannatta, K., Koontz, K., Kalinyak, K., Bukowski, W. M., & Davies, W. H. (1996). Peer relationships and emotional well-being of youngsters with sickle cell disease. *Child Development*, 67, 423–436.
- Noll, R. B., Zeller, M. H., Vannatta, K., Bukowski, W. M., & Davies, W. H. (1997). Potential bias in classroom research. *Journal of Clinical Child Psychology*, 26, 36–42.
- Perrin, E. C., Stein, R. E. K., & Drotar, D. (1991). Cautions in using the Child Behavior Checklist: Observations based on research about children with a chronic illness. *Journal of Pediatric Psychology*, 16, 411–421.
- Riekert, K. N., & Drotar, D. (1999). Who participates in research or adherence to treatment in insulin-dependent diabetes mellitus? Implications and recommendations for research. *Journal of Pediatric Psychology*, 24, 253–258.
- Silver, E. J., Westbrook, L. E., & Stein, R. E. K. (1995). Relationships of parental psychological distress to consequences of chronic health conditions in children. *Journal of Pediatric Psychology*, 23, 5–15.
- Sinois, P. A., Usner, D. W., Hill, S. D., Mitchell, W. G., Bab, J. F. Jr., Loveland, K. A., Loveland, K. A., Stehbens, J. A., Donfield, S. M., Maeder, M. A., Amodei, N., Constant, C. F., Nelson, M. D., & Willis, T. and the Hemophilia Growth and Development Study. (1998). *Journal of Pediatric Psychology*, 23, 45–56.
- Spiker, D., Kraemer, H. C., Scott, D. T., & Gross, R. T. (1991). Design issues in a randomized clinical trial of behavioral intervention: Insights from the Infant Health and Development Program. *Journal of Developmental and Behavioral Pediatrics*, 12, 386–393.
- Stein, R. E. K., Bauman, L. J., & Ireys, H. T. (1991). Who enrolls in prevention trials? Discordance in prevention of risk by professionals and participants. *American Journal of Community Psychology*, 19, 603–627.
- Thompson, R. J., Gil, K. M., Burbach, D. J., Keith, B. R., & Kinney, J. R. (1993). Role of maternal and child processes in the psychological adjustment of children with sickle cell disease. *Journal of Consulting and Clinical Psychology*, 61, 468–474.

- Verholst, F. C., & Koot, H. M. (1992). *Child psychiatric epidemiology. Concepts, methods and findings.* Newbury Park, CA: Sage.
- Wallander, J. L. (1992). Theory-driven research in pediatric psychology: A little bit on why and how. *Journal of Pediatric Psychology, 17*, 521–536.
- Wallander, J. L., Varni, J. W., Babani, L., Banis, H. T., DeHaan, C. B., & Wilcox, K. T. (1989). Disability parameters, chronic strains and adaptation of physically handicapped children and their mothers. *Journal of Pediatric Psychology, 4*, 23–42.
- Weinberger, S. A., Tublin, D. A., Ford, M. E., & Feldman, S. S. (1990). Preadolescent's social-emotional adjustment and selective attrition in family research. *Child Development, 61*, 1375–1380.
- Welch, W. P., Frank, R. G., & Costello, A. J. (1983). Missing data in psychiatric research: A solution. *Psychological Bulletin, 94*, 177–180.
- Widom, C. S. (1988). Sampling biases and implications for child abuse research. *American Journal of Orthopsychiatry, 58*, 260–270.

5

Strategies for Measurement and Psychological Assessment

**JAMES C. OVERHOLSER, ANTHONY SPIRITO,
and JULIA M. DIFILIPPO**

A number of special methodological issues and practical problems are involved in research with pediatric and child clinical populations. As in research with adults, there are generic difficulties involved in utilizing and developing measures that have both clinical utility and a solid research foundation. Researchers are also confronted with whether to select existing measures, to modify measures previously used for different purposes or populations, or to embark on the task of developing a new measure altogether. In addition, researchers must decide whether to use direct observation versus self-report measures. However, research with children and adolescents specifically raises a great many other formidable issues, including how to integrate information from multiple measures and/or multiple informants. The purpose of this chapter is to present a decision-making approach to a number of methodological and practical issues involved in measurement development with pediatric and child clinical populations. Whenever possible, we will provide the reader with examples from our own research to illustrate how we dealt with these issues.

For general texts on research methods in clinical psychology, researchers also may consult the following references, though they primarily focus on adults: Kazdin (1980, 1992), Bellack and Hersen (1984), and Creswell (1994). A related text in research with pediatric and clinical child populations is Achenbach (1978).

JAMES C. OVERHOLSER and JULIA M. DIFILIPPO • Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123. **ANTHONY SPIRITO** • Child and Family Psychiatry, Providence, Rhode Island 02903.

GENERAL ISSUES IN SELECTING PSYCHOLOGICAL MEASURES

Perhaps the most important consideration when selecting a measure is its potential relevance to the research question and design. Researchers must first conduct a thorough literature review of the psychological constructs under investigation and become familiar with the format, scoring, interpretation, and psychometric properties of existing measures. Researchers should examine individual items of the measure and its factors. Titles of measures may not accurately reflect their content. Careful attention should be paid to the theoretical foundation and original purpose of the measures, as well as the population for which the measures were designed. In addition, researchers should select measures that have a solid research base as well as practical utility for clinical settings. Some research measures are too cumbersome or time consuming to be used in most clinical settings. Conversely, some clinical measures may have face validity but lack the psychometric properties needed to trust them as reliable and valid measures of psychological constructs. Other relevant factors to consider are the cost of the measures themselves and the time and expenses involved in data collection and interpretation. Finally, when designing a study and choosing measures, researchers must address potential logistical challenges, such as how to obtain reports from all family members during an adolescent patient's brief stay at a psychiatric hospital unit.

TYPES OF PSYCHOLOGICAL MEASURES

Psychological assessment often relies on reports of typical performance (e.g., social skills measures) or psychological symptoms (e.g., depression severity). These measures provide useful information from the child or another knowledgeable informant. In addition, some psychological measures are considered tests of maximal performance (e.g., intelligence tests). Tests of maximal performance examine abilities when children are encouraged to perform at their best. Although still susceptible to a child's efforts to minimize strengths or exaggerate one's deficits ("faking bad"), tests of maximal performance are much less susceptible to attempts to exaggerate one's strengths ("faking good"). Finally, behavioral observations can be used to have a trained assistant observe the child in the natural environment in order to record specific aspects of the child's behavior. Although behavioral observations provide useful information, direct observation of children in their natural environment is too costly and cumbersome for most professionals conducting research or clinical assessment (Kerr, Hoier, & Versi, 1987). Because of the numerous difficulties encountered when obtaining observational data or information from the child's peers (e.g., siblings, friends, classmates), many clinical studies rely on the child's verbal self-report, whether obtained through written questionnaires or during a structured interview with the child or others familiar with the child.

Types of Psychological Measures

Self-report measures are the most common source of information used to evaluate problems such as childhood depression (Kerr et al., 1987). Self-report

measures provide an efficient means of gathering information about most types of psychological symptoms. Measures have been developed to quantify emotional distress (e.g., depression, anxiety, anger), behavioral dyscontrol (e.g., impulsivity), interpersonal problems (e.g., family conflict, peer socialization, social withdrawal), adaptive behaviors (e.g., coping skills), and treatment issues (e.g., compliance with treatment). These measures can be used to gather information relevant to the design, implementation, and evaluation of clinical services. Also, when evaluating emotional distress, researchers should consider that most self-report measures are more objective and direct in their assessment than are many projective measures (Brinkman, Overholser, & Klier, 1994).

Limitations of Self-Report Measures

Self-report measures have many limitations that can interfere with their ability to accurately assess psychological problems. Self-report measures require a certain level of reading skill and sophistication with paperwork and forms. Hence, younger children and those with cognitive impairments may not be capable of completing some questionnaires. Also, self-report measures rely on the subject's honest and accurate reporting of symptoms and events. Hence, self-report measures can be faked to exaggerate or minimize the severity of symptoms. In addition, it is not always reasonable to expect children to be able to describe or rate their subjective symptomatology. Young children may not be capable of adequately describing their symptoms of depression (Kazdin & Petti, 1982).

Most psychological problems are centered on the child's subjective experience. Hence, it becomes difficult to remove subjective biases from the measurement process. Sometimes, bias can be reduced by obtaining ratings from parents, teachers, or classmates. Peer rating measures can be useful in some research settings but are not widely used in clinical settings because of the time-consuming nature of the assessment procedures (Kazdin, 1981), as well as the need to obtain parental consent to participate in the assessment. In this section of the chapter, we will describe efforts we have undertaken to modify self-report measures, to develop short forms of existing measures, and to develop new measures in order to suit the research goals of the projects with pediatric and child clinical populations.

EXAMPLES OF SCALE DEVELOPMENT AND MODIFICATION

Applying an Established Measure with a New Population

In both research and clinical settings, a researcher may find a useful scale or interview that had been developed in a different setting or with a different population. For example, many established psychological measures have been developed on normal samples (e.g., college students or grade school children). When attempting to apply existing measures to other settings (e.g., medical hospital inpatients, psychiatric outpatients), the researcher must have evidence that the measure continues to assess the original construct and displays adequate psychometric properties (i.e., similar factor structure, internal consistency, interrater reliability, construct validity) relevant to the current use of the scale.

Researchers also must consider several dangers involved in making inferences from data that are based on a measure that was developed for use with a population that is different from that of the current study. For example, the Child Behavior Checklist (CBCL) (Achenbach, 1991) is one of the most widely used measures in research on child psychopathology. Given that the CBCL was designed to differentiate between children who were referred for mental health treatment and those who were not referred, the measure may be much less effective in making distinctions between children with mild, subclinical symptoms (Drotar, Stein, & Perrin, 1995). In addition, if researchers decide to use the CBCL with a chronic illness population, they must decide how use of the measure will affect their study. For example, the CBCL subscale that includes somatization items should not be reported. The social competence scale and the items pertaining to physical symptoms on the CBCL may not be valid for special populations, such as children with chronic physical illness, because chronic illness may limit opportunities for participation in certain social activities (Drotar et al., 1995). Therefore, a low *T*-score on the social competence scale does not necessarily mean children with chronic illness have poorer social adjustment. Consequently, a control group will be mandatory in a design using the CBCL with chronically ill children, and the scale scores will have to be calculated in such a way as to not bias the scores of the chronic illness group. For instance, the somatic items may need to be dropped and the chronic illness and control group could be compared on raw scores rather than *T*-scores. Another alternative would be to score the somatic items as zero for the chronic illness group and to calculate *T*-scores after this adjustment. The modifications mentioned above are examples designed to ensure that data obtained from an established measure accurately reflects the new population's functioning.

In our own work, we attempted to validate a suicide lethality rating scale that originally had been designed for adult suicide attempters (Spirito, Brown, Overholser, Fritz, & Bond, 1991a). Our study hoped to determine whether suicide lethality ratings could be used with adolescents who had recently attempted suicide. The Risk-Rescue Rating Scale (Weissman & Worden, 1972) was developed to quantify variables relevant to a suicide attempt after a trained interviewer conducted a semi-structured interview with the patient. The Risk-Rescue Rating Scale was used to evaluate suicide risk in 109 adolescents. Research assistants were instructed in scoring the scale after a series of training steps designed to maximize chances for obtaining reliability. Additional decision criteria were added to the scale to help in this process.

We examined the interrater reliability of the Risk-Rescue Rating Scale with Cohen's kappa, a conservative estimate of reliability. Despite the training efforts, Cohen's kappa interrater reliability was rather low but considered adequate ($k = .71$). Upon closer inspection, Cohen's kappa may have been low because of the high base rate for many items. If the scores on items of the Risk-Rescue Rating Scale had been more evenly distributed, disagreements would not have so negatively affected the kappa coefficient. We could have chosen a less stringent statistic (e.g., percent agreement), given the nature of the score distribution. However, the problems with Cohen's kappa were most instructive in determining the clinical utility of the scale. If the variability of scores is limited, the Risk-Rescue Rating Scale is not likely to be useful in predicting or in discriminating different subgroups of suicide attempters. Thus, in our study, the choice of the statistic helped elucidate the limitations of the scale for the population under investigation.

Based on the data described above, we abandoned the use of the Risk–Rescue Rating Scale in further studies with adolescent suicide attempters. Instead, we systematically began to investigate the usefulness of the Suicide Intent Scale (SIS) (Beck, Schuyler, & Herman, 1974) with adolescent suicide attempters. We chose the SIS because the measure was designed to assess both subjective and objective aspects of intent, unlike the Risk–Rescue Rating Scale, which only assesses objective aspects of an attempt. The subjective aspects of intent refer to items such as ambivalence about living and expectation of the fatality of the attempt. Objective aspects include items such as precautions against discovery and acting to gain help after the attempt. Thus, the SIS appeared to have more to offer clinicians than the risk–rescue because the SIS contains sufficient objective information about the attempt as well as items regarding the intent of the adolescent's suicidal behavior. Despite the more comprehensive nature of the scale, the SIS had been used rarely with adolescents. In our study (Spirito, Sterling, Donaldson, & Arrigan, 1996), 190 adolescent suicide attempters were rated on the SIS by either a psychologist or psychiatrist. Internal consistency was demonstrated for the entire scale and the subjective subscale but not for the objective subscale. Three factors emerged from a factor analysis, with most of the subjective items loading on one factor and four out of seven objective items loading on a second factor. The third factor was composed of items related to planning the attempt. Thus, the subjective–objective theoretical grouping of items on the SIS, which has been confirmed by several factor analyses of the scale with adults, did not fit as well with adolescent suicide attempters. That is, the results of the factor analysis indicated that the planning of the attempt is a separate factor to consider in understanding an adolescent's suicide attempt. Based on our research, the SIS is potentially useful to clinicians evaluating adolescent suicide attempters if the factor scores specific to adolescents are applied rather than the adult-derived subscales or total score. In addition, the factor scores were found to be differentially correlated with measures of depression and hopelessness, suggesting that the factors tap different aspects of suicide intent and might be used in future studies to help predict follow-up and treatment outcome.

Modifying an Existing Self-Report Scale

Researchers in clinical child and pediatric psychology are often faced with the problem of modifying an existing scale. For example, there may be questions about the applicability of measures to the population of interest (e.g., children of different cultural and ethnic groups, children with chronic physical illness). In addition, an existing measure may be limited in content or lack clinical relevance.

In a study conducted by our research group (Lehnert, Overholser, & Adams, 1996), we grappled with the limits of self-report measures. We were interested in examining cognitive biases that may play a role in adolescent depression, but we felt that most self-report measures failed to assess the client's actual cognitive content. Instead, many measures that purport to rate cognitive biases seem to provide an estimate of depression severity. Also, many measures of cognitive biases have been developed on college student samples and may be phrased in terminology or sentence structure that would be inappropriate for use with children or adolescents.

In our study, two groups of adolescents (115 adolescent psychiatric inpatients and 102 high school students) completed the assessment. We examined how re-

searchers can measure cognitions, a question that has great relevance for cognitive therapy. Our study used the Incomplete Sentence Blank (Rotter, 1950), an assessment method that is commonly used in clinical settings, and we developed a new scoring system so that cognitive content could be scored in a reliable manner. We named our modified measure the Cognition Rating Form, which uses incomplete sentence stems to allow the adolescent a semistructured stimulus to record their thoughts and reactions. For example, the sentence stem "I hate ..." can be completed in a variety of ways. Instead of forcing subjects to agree or disagree with a predetermined list of statements, the Cognition Rating Form allows subjects to answer each item with an unlimited range of possible answers. Thus, the Cognition Rating Form assesses cognitive patterns in a more naturalistic way than allowed by most other measures. Despite the open-ended nature of the assessment, we developed scoring procedures that can quickly derive numerical ratings for each subject. After the adolescents completed 25 sentences, their responses were scored by two trained graduate students. Each sentence was scored for the presence or absence of several types of cognitive biases, including a negative view of self, negative view of the world, negative view of the future, passive thinking, self-blaming, and other-blaming. In order to later validate the modified measure, we also asked subjects to complete standard, well-established measures of depression, hopelessness, and negative cognitions. Depression and hopelessness have been shown to be related to negative cognitions, and the established measure of negative cognitions could be used to validate our modified measure.

Reliability of the Cognition Rating Form was examined by assessing interrater agreement and internal consistency. We found adequate interrater agreement based on a criterion of $r = .70$ or greater for the cognitions on the measure. In addition, we determined internal consistency using a statistical technique known as a principal components factor analysis. Principal components factor analysis summarizes the patterns of correlations among items and reduces the measure to a few homogeneous subscales. The principal components factor analysis on the Cognition Rating Form found that the measure consisted of four factors, together accounting for 73.9% of the total variance. We named the factors Cognitive Complexity, Negative Cognitions, Negative Others, and Positive View of the Future.

When validating a new scale, it is important to examine its performance relative to an existing measure of the same construct (e.g., negative cognitions), as well as theoretically related constructs (e.g., depression and hopelessness). There may be times when the relationship between the new scale and an existing scale may not be very robust because the new scale measures a slightly different aspect of the construct, the method varies somewhat (i.e., forced-choice versus open-ended questions), or the new measure is better than the existing measure. In order to validate the Cognition Rating Form, we correlated the Cognition Rating Form with established measures of depression, hopelessness, and negative cognitions, which theoretically should tap into the same construct as our scale. Findings that the Cognition Rating Form correlated in the anticipated directions with standard measures of depression, hopelessness, and negative cognitions provided evidence of concurrent validity. Furthermore, through two sets of parallel discriminant function analyses, we examined how well the Cognition Rating Form, compared with other established measures of depression, could correctly classify depressed adolescent psychiatric inpatients versus high school controls. We found that the

Cognition Rating Form yielded comparable, if not better, classification of adolescents, relative to an established measure of depression severity.

Although we demonstrated adequate reliability and validity for the Cognition Rating Form, our work in modifying an existing scale is far from over. For example, additional studies are needed to examine other cognitive categories that should be included in future revisions of the measure. Furthermore, we may consider how to measure cognitions specific to other forms of psychopathology, such as anxiety. Finally, future studies are recommended to examine the temporal stability and predictive validity of the Cognition Rating Form.

Developing a Short-Form of an Existing Interview Measure

Researchers of pediatric and child clinical populations often need measures that are brief and focused. Although a particular measure may be well established and theoretically sound, it can present significant practical difficulties when it is too time consuming. For this reason, researchers may choose to develop a short form of an existing measure. Although previous studies may have demonstrated adequate psychometric properties of the existing measure, a short-form version must retain reliability and validity.

In many of our studies, we have used the Children's Depression Rating Scale (CDRS) (Poznanski, Cook, & Carroll, 1979) as a useful measure of depression severity. The CDRS, a child version of the widely used Hamilton Rating Scale for Depression (Hamilton, 1960), is a structured interview that evaluates the presence and severity of 17 common symptoms of depression. A trained interviewer conducts a naturalistic, semistructured interview to gather detailed information about a range of depressive symptoms. The CDRS provides qualitative and quantitative information that is useful in both research and clinical assessment. However, the CDRS is fairly long, often requiring 45 minutes to complete the interview.

In our study (Overholser, Brinkman, Lehnert, & Ricciardi, 1995), we evaluated a total of 228 adolescent psychiatric inpatients using the CDRS and other measures relevant to depression. We developed a short form of the CDRS that included the five items we felt best captured the nature of depression for most adolescents. The items we included in the short form were dysphoric mood, anhedonia, social withdrawal, low self-esteem, and fatigue. We tried to select items that evaluated different domains of depressive symptomatology (e.g., affective, behavioral, cognitive, and somatic manifestations of depression). Results showed that little information was lost when switching from the long form to the short form of the CDRS. The two versions correlated highly ($r = .91$) and were related to other measures of depression in a comparable manner. This new short form can be used in both research and clinical settings to provide an efficient, focused, interview-based screening measure of depression severity. By itself, it is inadequate for making a diagnosis of major depression, but it can help to identify those individuals who would benefit from a more thorough evaluation of their depression.

Developing a New Self-Report Measure: Difficulties in Validation

In both research and clinical settings, researchers often must decide whether to use existing measures or attempt to create a new measure particularly suited to

their current needs. Developing a new measure can become an exciting project whereby a new investigator has an opportunity to develop a scale of a psychological construct that has not been adequately studied in prior research. Sometimes, a new measure can create a wave of excitement in an area, stimulating further research by other investigators. However, developing a new measure requires an extensive commitment of time and energy. One study is never enough to document the psychometric properties of a new scale. Furthermore, many investigators are reluctant to adopt a new measure unless there are several published studies examining its psychometric properties. Hence, an investigator who decides to develop a new measure must dedicate several years of work evaluating and revising the scale.

The majority of our test development efforts have resulted from our attempts to assess the central factors we considered to be important aspects of some clinical problem. Our interest in impulsivity arose from our research and therapy with adolescent suicide attempters. Impulsive suicide attempts are common during adolescence. In the course of trying to discriminate subgroups of adolescent suicide attempters, a very heterogeneous population, we found ourselves looking for a measure of impulsivity. Although we were able to classify impulsive suicide attempts (Brown, Overholser, Spirito, & Fritz, 1991), we were interested in looking at the contribution of a more general impulsive behavioral style to suicidal behavior.

We examined the published literature and did not find an existing measure that would adequately meet our needs. There have been a number of measures designed to quantify impulsive tendencies with adolescents, but when examined closely many of the established measures contain items tapping cognitive impulsivity, aggressiveness, risk-taking, or delinquent behaviors. Although these latter items are strongly related to impulsive behavior, we were trying to thoroughly describe risk-taking behaviors and aggressiveness with separate measures. We sought to develop a measure that would focus specifically on impulsivity. In order to accomplish this task, we first reviewed the literature and collected scales that examined impulsivity, aggressiveness, and risk-taking behaviors. We also used the criteria for impulsivity as defined by the *Diagnostic and Statistical Manual of Mental Disorders*, 3rd edition, revised (DSM-III-R) (American Psychiatric Association, 1987). Through this search, we were able to identify a number of items from these different scales that seemed to focus specifically on impulsive behavior rather than cognitive impulsivity. Most of the items were adapted from other scales but we also developed a few new items. These items were selected to assess impulsive behavior and not be confounded by either anger or risk-taking behaviors that are found in many of the other scales of impulsivity. Some examples of items in the new impulsivity measure include: "I am impatient and have a hard time waiting for things" and "I do things that I later wish I hadn't."

In order to develop the scale, we first arranged to have the scale [the Impulsive Behavior Checklist-(IBC)] administered to 167 high school students. We collected both reliability and validity data with this one sample. At baseline, we administered our new impulsivity checklist along with the Impulse Control subscales from both the Millon Adolescent Personality Inventory (MAPI) (Millon, Green, & Meagher, 1982) and the Offer Self-Image Questionnaire (OSIQ) (Offer, Ostrov, & Howard, 1982), as well as the Anger Expression Scale from the State-Trait Anger Scale (Spielberger, Jacobs, Russell, & Crane, 1983). At baseline, students were administered the IBC and the OSIQ Impulse Control Subscale. Four weeks later, they were

asked to complete the IBC once again as well as the Anger Expression Scale and the MAPI Impulse Control Scale.

Internal consistency was evaluated by examining item-to-total correlations at both pretest and posttest. The average item-total correlations were in the moderate range for both pre- and posttest scores. A split-half correlation of 0.55 was converted to an estimated full-length split-half correlation of 0.70 using the Spearman-Brown formula. A Cronbach's alpha of .65 was attained as an estimate of the lower limit on reliability. Test-retest reliability was somewhat higher, with a correlation of .77. Thus, the scale demonstrated fair homogeneity and relatively good temporal stability.

Concurrent validity findings were somewhat more concerning. Concurrent validity examines the relationship between different measures of the same construct, all collected at the same time. Most of the correlations between scales ranged from 0.38 to 0.44 for the OSIQ and the MAPI Impulse Control Scales. Correlations with the anger scales ranged from .27 to .54. Because the scales from the MAPI and the OSIQ were labeled impulse control but also contained items focused on anger and other issues, these correlations of 0.3 to 0.4 were acceptable. However, analyses did not provide evidence for the concurrent validity of the scale. In fact, when the sample was split into impulsive versus nonimpulsive subjects by a median split on the IBC, the impulsive group had a lower score on the OSIQ Impulse Control Scale than the nonimpulsive group, while the opposite was true for the MAPI Impulse Control Scale. It was clear that to be able to establish validity, we would need other ratings of behavior. We considered using parent and teacher ratings of impulsivity. Another possibility involved rating behavioral indicators of impulsive responding, motoric restlessness, and degree of persistence while working on a structured task. Using neuropsychological tasks, such as the Stroop, Trail-Making, or the Continuous Performance Test, was another approach we considered.

Although our items had face validity and were likely to be useful to clinicians, the cost-benefit ratio of investing the time to validate the scale argued against putting much further effort into the development of this measure. The type of study needed to develop a clear assessment of impulsivity was published by White and colleagues (1994). This multimethod, multisource assessment of impulsivity revealed two impulsivity factors, one cognitive and one behavioral. The items that loaded on the behavioral impulsivity factor included an observer rating of motor restlessness, teacher-reported impulsivity, self-reported impulsivity on the Eysenck (Eysenck Impulsiveness Scale) (Eysenck et al., 1989), observer-rated impatience-impersistence, and a Q-sort for parents regarding undercontrol. The comprehensiveness of the White et al. (1994) study underscores the fact that researchers who propose a new scale should assume that scale development efforts will become a major focus of their work for at least 5 years.

Problems in the Development of Brief Measures: Reliability and Factor Analysis

Researchers developing new measures also face the challenge of demonstrating adequate reliability. As part of ongoing work in examining adjustment in pediatric populations, we initiated the development of the Kidcope (Spirito, Stark, & Williams, 1988). At the time, we were interested in moving beyond an assessment

of psychopathology and beginning to evaluate the adaptive mechanisms used by chronically ill children. Coping was one of several areas of adaptive (e.g., social adjustment) and maladaptive functioning (e.g., depression and anxiety) that we planned on measuring. Thus, we needed a relatively brief measure. When we began this project in 1986, there were no self-report measures in the literature that had been used routinely to examine coping in children. Several interviews had been developed to assess coping, but these were rather detailed and would not be feasible given the length of the proposed battery. Thus, we set out to develop our own measure of coping. In order to do so, we initially selected a pool of 24 items that appeared to tap the coping strategies most commonly found in various studies on coping with adults as well as those that would tap the coping strategies we were most interested in studying in a pediatric chronic illness population. The initial item pool was purposely small in the hope that we could develop a brief questionnaire.

We administered the scale to high school students and then factor analyzed the Kidcope scores. However, as is true in many factor analyses, some factors were not as conceptually meaningful as we would have hoped (i.e., the coping strategy "blame" loaded with "wishful thinking"). Thirteen items loaded on six factors reflecting a wide range of cognitive and behavioral coping strategies. Thus, rather than continue to develop a longer scale, we decided to move more toward developing a screening measure with a broad range of cognitive and behavioral coping strategies.

With only one item tapping each of the ten coping strategies, documenting the psychometric properties of such a checklist was much more problematic. Nonetheless, we proceeded to conduct both reliability and validity studies of the measure with some success (Spirito et al., 1988; Spirito, Stark, Grace, & Stamoulis, 1991). The brevity of the Kidcope was especially problematic in establishing its reliability. We attempted to calculate reliability item per item using test-retest reliability coefficients, with test-retest correlations on most items in the moderate range. We were unable to calculate internal consistency coefficients as another marker of reliability because each of the ten items on the scale was designed to be conceptually distinct, so that we would not expect there to be internal consistency among the items.

Many researchers have been interested in analyzing the ten coping strategies of the Kidcope using a higher-order factor structure (e.g., positive-negative, approach-avoidance), rather than using the individual coping strategies. Summary scores are particularly useful in regression analyses when the authors are interested in examining the contribution of coping as one of several variables related to overall adjustment. However, the design of the Kidcope does not lend itself well to factor analysis. We have attempted to factor analyze the scale using a large sample of subjects (approximately 3000) as well as with smaller, more defined subpopulations. Although the factor structure tends to be more distinct when we are using subpopulations, it changes depending on the situation to which the subjects are responding. The changing factor structure suggests that categorization of coping strategies cannot be separated from the situation. Hence, the function of a coping strategy does not remain identical across situations. For example, distraction may be an adaptive alternative in one situation and maladaptive in another situation. Thus, there are no easy answers to the problem of higher-order structure. Several strategies might be implemented to deal with this problem (Spirito, 1996). One

approach is to conduct a factor analysis for each data set. This requires between 150 and 200 subjects, so it is not often feasible in pediatric psychology where studies are often conducted with a small number of subjects. A second approach is to gather experts who are presented with the situation in which coping is necessary. The experts then decide how to group the ten different strategies based on some conceptual model applied in the situation.

Educating Other Researchers About Your New or Revised Measure

After developing a new or revised measure and demonstrating its psychometric properties, it is important for the researcher to educate others regarding its application, strengths, and weaknesses. A cover letter delineating the best applications of the measure as well as potential problems is often helpful. For example, in the case of the Kidcope mentioned in the previous section, we explain that the Kidcope is a brief screening measure with limited higher-order factor structure. When coping is the major variable of interest, we believe that there is no substitute for an interview. Despite the time and effort the researchers have put into developing a new measure, we advise that they clearly state its weaknesses and limitations. Otherwise, the new measure may disappoint other researchers who find that the data collected do not meet their needs after the project has been completed.

INTEGRATING MULTIPLE MEASURES AND MULTIPLE INFORMANTS

Whenever possible, a comprehensive assessment should incorporate the use of empirically established scales that gather information on several domains of functioning from multiple sources (Achenbach, 1993). Although it is preferable to use multiple measures and multiple informants, the interpretation of multiple types of data can be complicated. Below, we outline some of the problems encountered and ways to address potential problems.

Combining Information from Multiple Measures

When multiple measures are used, combining data from the different measures can be complicated. For example, when two self-report scales are obtained from the same subject, a high degree of consistency has been obtained in the assessment of depression (Kazdin, French, Unis, & Esveldt-Dawson, 1983). Even so, the relationship of these self-report measures to other variables in a study may differ slightly, or the number of subjects who completed each measure may be relatively small. Consequently, in some studies, researchers might want to combine the scales in some way.

In one of our studies (Jelalian et al., 1997), we used three brief, established self-report measures of varying lengths (6, 4, and 2 items) to assess adolescent risk-taking tendencies. Although each scale revealed comparable results, the relationship of each scale to the other measures in our battery was slightly different. Because we had a large sample (1400 high school students), we decided to combine the items and conduct a factor analysis. A single factor emerged that consisted of ten of the 12 items and accounted for 31% of the variance. The alpha coefficient of

.81 was better than any of the individual scales, probably due to the better psychometric properties of a longer scale. Conducting a factor analysis simplified later analyses and provided sufficient information about the construct for the purposes of our study.

A multimethod assessment often results in discrepancies across the different types of measures. As noted by Kazdin and Pettit (1982), one measure may emphasize certain qualities (e.g., somatic aspects of depression), while another measure may focus on different aspects of the same construct (e.g., cognitive aspects of depression). Also, different types of measures (e.g., self-report scales vs. projective techniques) are devised to measure different psychological qualities. For example, the Rorschach Inkblot Test can be a useful measure of general personality style, but it has not been found to be accurate in its ability to quantify depression (Lipovsky, Finch, & Belter, 1989). The investigator must determine which measures are best suited for the assessment of specific psychological constructs. Simply because a scale purports to measure a particular quality does not guarantee its validity.

Combining Information from Multiple Informants

A comprehensive assessment of behavioral and emotional functioning among children and adolescents relies on information from a variety of people, including parents, teachers, peers, and the children themselves. Researchers and clinicians who obtain data from multiple informants are provided with a more complete picture of how the child's functioning varies from one context to another and how informants differ with respect to their sensitivity to and tolerance for problematic child behavior (Achenbach, 1991). Multiple informants are highly desirable for assessing young children who may not have the cognitive ability or psychological insight to express the range, severity, and duration of their symptoms as well as the impact of their problems on others. Similarly, school-aged children and adolescents may fail to report disruptive behaviors, such as hyperactivity, because they may not consider them to be problematic.

Factors Associated with Discrepant Reports among Multiple Informants

When data from multiple informants converge on a particular disorder, researchers and clinicians can feel more confident that their findings are valid and reliable. However, numerous studies indicate low-to-moderate correlations between reports from multiple informants, even when considering the same aspect of the child's behavioral-emotional functioning. For example, little agreement is obtained when depression scores are gathered from child, peers, and teacher (e.g., Crowley, Worchel, & Ash, 1992).

Based on a meta-analysis of 269 samples in 119 studies involving ratings by multiple informants, Achenbach, McConaughy, and Howell (1987) concluded that average Pearson correlations differ, depending on which informants are compared. For example, mean Pearson correlations were .60 between similar informants (e.g., mother and father), .28 between different types of informants (e.g., parents and teacher), and .22 between the child and other informants. Typically, mothers and fathers agree on ratings of depression in the child, while neither parent shows

much agreement with the child's report of depression severity (Kazdin et al., 1983). Fathers often display poor agreement with their children on the frequency of depressive symptoms (Ivens & Rehm, 1988). However, family members can provide information about factors other than psychiatric symptoms, such as conflict within the home. Maladaptive family functioning can be related to depression and suicidal tendencies among adolescents (Adams, Overholser, & Lehnert, 1994).

In light of low correlations between different types of informants, different informants contribute a considerable amount of unique variance. Children provide more informative reports of depressive symptomatology than do mothers; however, mothers demonstrate better recall for their children's depression at 2-year follow-up (Fendrich, Weissman, & Warner, 1991). In other words, both child and mother reports contribute valuable but different information to the assessment process. In addition, the mean correlation of .22 between ratings by children and other informants indicates that children's self-reports cannot substitute for reports by others (Achenbach et al., 1987).

In general, correlations between informant reports are often higher for 6- to 11-year-olds than for adolescents (Achenbach et al., 1987). Adolescents may be more likely to withhold information from their parents and participate in activities outside the home that parents do not supervise. Thus, like adults, adolescents should be interviewed to provide their subjective accounts of their own psychological functioning (Tarullo, Richardson, Radke-Yarrow, & Martinez, 1995; Verhulst & van der Ende, 1992). In addition, peer reports may be of increased utility during adolescence because adolescents may confide in peers rather than parents.

Correlations between informants differ depending on the aspect of the child's psychological functioning under investigation (Achenbach et al., 1987). Stronger correlations between informants are found when assessing factors that can be directly observed (such as undercontrolled conduct problems and hyperactivity) instead of qualities that must be inferred (such as depression, anxiety, and obsessions) (Edelbrock, Costello, Dulcan, Calabro Conover, & Kala, 1986; Huddleston & Rust, 1994). Inter-rater agreement also has been found to be higher regarding observable behaviors that violate social role expectations (e.g., conduct problems among girls) (Jensen, Traylor, Xenakis, & Davis, 1988a; Tarullo et al., 1995).

Children tend to report more subjective symptoms and covert behavior, while adults report more behavioral symptoms (Herjanic & Reich, 1982). Parents are more likely to express concern about child behaviors that they find disturbing, such as disobedience, loudness, inattention, and overactivity. In contrast, children and adolescents may be more distressed by difficulties with feelings (Edelbrock et al., 1986).

Although the majority of parent reports have traditionally involved mothers, some recent studies have examined mother's versus father's perceptions of child psychological functioning (Jensen et al., 1988a; Webster-Stratton, 1988). Paternal psychopathology can influence perceptions of daughters more than sons, whereas mothers' symptoms affect reports of sons more than daughters (Jensen et al., 1988a). Despite claims that depressed mothers provide distorted and inflated reports about their children's problems, a review of 22 studies concluded that research has not been adequately designed to directly test such assumptions (Richters, 1992). Family status (divorced versus intact), sex of the parent and child, life stressors, sibling

position, and familiarity of the child to the rater are other factors that should be considered when conducting multiple informant assessments (Jensen, Xenakis, Davis, & Degroot, 1988b).

Utility of Discrepant Reports from Multiple Informants

Discrepancies among informants' reports may provide information regarding how the child's behavior varies across situations (Achenbach et al., 1987). Discrepant reports also can indicate that clinical interventions may require specific targets in a particular context. For example, discrepant findings may occur when a teacher reports that the child's activity level is within the average range, while parents report that the child is quite hyperactive at home. Based on the teacher-parent discrepancy, a clinical intervention might be proposed that focuses on decreasing the child's inattention specifically at home, using parent training and behavioral techniques. The clinician also might consider altering the parents' perceptions of the child, thus, increasing the parents' tolerance for some degree of child inattention.

Although discrepancies between informants may become confusing, obtaining data from several sources has been found to increase diagnostic accuracy. For example, Drotar, Agle, Eckl, and Thompson (1996) emphasize the importance of considering information from mothers of highly defensive children with chronic illness (hemophilia) when evaluating the children's psychological distress. Defensive children are defined as those who are motivated to report low levels of psychological distress (e.g., depression, anxiety) despite physiological evidence and other informants' consistent reports that the children are in fact highly distressed (Weinberger & Schwartz, 1990; Weinberger, 1996). Thus, sole reliance on children's self-report runs the risk of providing an underestimation of psychological distress among highly defensive individuals, and more accurate information may be obtained from multiple sources (Drotar et al., 1996; Weinberger, 1996). In addition, accurate classification of major diagnostic groups of externalizing, internalizing, and mixed disorders among outpatient boys has been found to increase significantly when ratings from both parents, rather than one parent, are combined with teacher report (Mattison, Bagnato, & Strickler, 1987). Despite potential difficulties integrating discrepant information, obtaining data from multiple sources has some clear advantages in terms of providing a more complete assessment of the child (Reich & Earls, 1987).

Approaches to the Integration of Discrepant Data from Multiple Informants

Integrating discrepant data from multiple informants is one of the most challenging tasks of child psychological assessment. In one of our studies (Silverman & Overholser, 1995), we examined reports of depression and family problems as described by 60 adolescents and their parents. On measures of family problems [e.g., communication, control, and role performance subscales from the Family Assessment Measure (Skinner, Steinhauer, & Santa-Barbara, 1990)], moderate correlations were obtained between adolescents and both parents. However, between-group comparisons (i.e., paired *t*-tests) revealed significant differences between information obtained from child versus parent informants on many measures of

family conflict. Thus, we were left with the dilemma of whether to view the different informants as agreeing or disagreeing. There are no clear guidelines for whose report should be trusted when discrepant information is obtained. Because previous research (e.g., Brinkman et al., 1994) has found that direct measures can adequately assess emotional distress in adolescents, it is reasonable to assume that most adolescent patients do not deny or minimize their level of distress. However, ratings of emotional distress are different from views of family functioning in that two informants can report different perceptions of the family, and both can be right.

Achenbach (1991) presents some general recommendations for integrating data from multiple sources via a taxonomic decision tree: The first question to ask is whether any of the informants' reports are within the clinical range. If the answer is yes, the researcher should identify the number of problem areas and whether all sources report the same syndromes. If the same syndromes are reported by all sources, the researcher may conclude that the child has multiple problems manifested in several contexts. However, if some sources do not report significant problems in a particular area, the researcher must then determine whether the child's behavior actually differs much among contexts. Based on this information, the decision can be made to either (1) target different behaviors for change in different contexts, or (2) alter the perceptions of some informants (Achenbach, 1991).

Two other major approaches to aggregation of information from multiple sources include simple information-combining schemes and complex information-combining schemes. Advantages and disadvantages of these approaches will be discussed.

Simple Information-Combining Schemes

Simple information-combining schemes assume that symptoms are either present or absent and that all informants provide data of comparable utility (Bird, Gould, & Staghezza, 1992). Thus, all information is weighted equally. Simple schemes ignore discrepancies and accept a symptom or criterion as present, as long as at least one informant endorses the particular item. For example, oppositional-defiant behavior would be considered present in an adolescent, even if only the father reported characteristic symptoms. Although simple information-combining schemes are quite easy to use, they assume that informants are equally reliable and they ignore discrepancies in child behavioral and emotional functioning in different contexts. In addition, simple schemes may produce high false-positive rates, thus, overestimating child psychopathology.

Complex Information-Combining Schemes

Complex information-combining schemes place greater weight on data provided by an "optimal" or preferred informant; the optimal informant depends on the symptoms under investigation (Loeber, Green, Lahey, & Stouthamer-Loeber, 1991). As discussed earlier, there is evidence that children are better informants of internalizing symptoms, while parents provide more reliable information regarding children's externalizing symptoms. Teachers may be the preferable informant of

the child's learning difficulties. For these reasons, researchers and clinicians might weigh child reports of depression more heavily than parent reports and they could weigh parent reports of hyperactivity more heavily than child reports.

Although the rationale for complex information-combining schemes is appealing, operationalizing specific decision rules about weighing information can be quite challenging (Bird et al., 1992; Piacentini, Cohen, & Cohen, 1992). Two main statistical approaches can be used in order to determine an "optimal" informant. One complex scheme involves logistic regression analysis, with the person's report most predictive of the clinician's diagnosis serving as the "optimal" informant. For each diagnosis, the researchers determine whether the parents' or child's responses were most predictive of the clinical diagnosis (outcome variable). In terms of the logistic regression, the optimal informant is that which provides the highest regression coefficient (Bird et al., 1992). However, if the parents' and child's responses are both significantly related to the diagnosis and there is no statistically significant difference in their regression coefficients, the sources are considered interchangeable and are both included in the regression. Another complex scheme for selecting the "optimal" informant is the conditional probability approach (Loeber, Green, Lahey, & Stouthamer-Loeber, 1989). Conditional agreement is the probability that the child will report the presence of a problem, given that the parents also reported the same problem, and vice versa. The conditional probability approach involves the percentage of agreement on endorsed problems. The informant with the higher conditional agreement is considered to be the more "necessary" informant. When conditional agreement for both or several informants is very high, informants are considered interchangeable. Thus, it would not be necessary for the researcher or clinician to obtain data from both informants regarding that particular issue. In situations involving low conditional probability for both informants, multiple informants are necessary (Loeber et al., 1989).

Comparisons of simple schemes versus the two methods of complex schemes for integrating information have found that no method offers a distinct advantage over another (Bird et al., 1992; Piacentini et al., 1992). Therefore, complex schemes may be difficult for the researcher to use and do not appear to provide significantly more accurate information than simple schemes.

A recent comparison of methods used to integrate assessment data from multiple informants (Offord et al., 1996) demonstrates the disadvantages of combining data from informants by simple or complex schemes. By keeping data from multiple informants distinct and by conceptualizing child psychiatric disorders as informant-specific, researchers may investigate unique patterns of associated features, specific etiology, and treatment response (Offord et al., 1996). For example, Offord and colleagues (1996) found that children identified by their *parents* as having conduct disorder are more likely to have a depressed parent and to live in a dysfunctional family, while children identified by their *teachers* as having conduct disorder are more likely to be males living in families of low socioeconomic status. Thus, based on the comparison of integration methods, Offord and colleagues (1996) recommend that researchers keep data separate by informant rather than utilizing simple or complex schemes. Researchers still have the option of combining the information from different informants at a later time without losing valuable information regarding the sources of the data.

Practical Difficulties Using Multiple Measures and Informants

Interpreting data from multiple sources can be complex, but obtaining the data itself is fraught with difficulties. In several of our research projects, we had the cooperation of the hospital admissions office when requesting informed consent to participate in the study. At the time the adolescent was admitted to the hospital, the admissions staff gave the parents a packet for our research. The packet included a brief explanation of the study, an informed consent sheet, and several questionnaires. Parents were asked to sign the consent form, complete all questionnaires, and return the packet to the admissions office within 3 days. We collected packets from the admissions office and then contacted the adolescent to request their participation in the study. The above process was slow and cumbersome and there were many points where information could get lost. Sometimes, the parents had completed the consent form and parent measures in a timely manner but the packets were not available to us until several days after the adolescent had been discharged from the hospital. These potential subjects were lost from the study. Therefore, researchers planning to obtain data from multiple informants must anticipate difficulties in gathering and interpreting such information.

FINAL RECOMMENDATIONS

Designing a Battery

When designing an assessment battery for clinical research, the investigator should keep several guidelines in mind. Measures should be selected based on the published literature as well as their relevance to clinical care. Numerous measures of psychological functioning have been published. Researchers should be cautious in deciding which scale to use, and they should rely on the empirical evidence that supports the reliability and validity of the scale. When no measure has been developed to suit the needs of a particular study, it may become necessary to design a study around the development of a new or modified measure.

Despite the value of a comprehensive assessment battery, the assessment protocol should be kept to a reasonable length. A manageable amount of time for children or adolescents to sit and answer questions is usually limited to 1 hour of data collection per subject. The assessment battery can be kept brief and focused by selecting measures that are relevant to clinical care and investigate one research question at a time. Thus, studying one area (e.g., depression severity) in depth can allow the investigator to examine the differential utility of several measures of depression. Also, an assessment battery works best when it includes a mixture of new and established measures. The new measures may provide information about the separate but related constructs, while the established measures provide a base for comparison purposes. However, because of the difficulties involved in establishing the reliability and validity of a new scale, it is often best to use an existing scale with few modifications to its established procedures.

Prior to beginning the study, the assessment battery should be evaluated for its adequacy in answering the central questions of the study. The assessment battery should be focused directly on the specific research questions under examination.

When a project becomes too broad in scope, measures may be collected but never analyzed, an unacceptable circumstance for both the research team and the subjects. Whenever possible, longitudinal designs should be considered, even if it is only short term (e.g., 3–6 weeks) and uses only a few assessment points (e.g., pre- and posttreatment evaluations).

When conducting research on clinical populations, researchers should consider that children may not complete measures accurately and thoroughly. Some children may have difficulties reading standardized questionnaires, may omit items, or may become bored and restless halfway through the assessment. We have found it most useful to have a trained research assistant read the questions to each child individually in order to ensure that all questions are understood and answered properly. If the assessment packet is long, the test administrator can judge how much to accomplish in one sitting. Whenever a child speaks English as a second language, it is essential to have a translator available to ensure proper understanding and contextual interpretation of the questions.

Training Research Assistants

Investigators should ensure the data are collected properly. It is essential to have well-trained research assistants help with the collection, scoring, and coding of data. Prior to beginning with data collection, all research assistants should be trained in the basics of standardized testing and should be given preparatory training about the particular problem area under investigation. The data collection processes should be evaluated on a regular basis to ensure that all members of the research team are following the same procedures. When it has been determined that some data were collected inappropriately or subjects responded randomly, the data should be omitted from the study.

Managing Data

Data should be entered into a computer database on a regular basis and periodically summarized. Inspection of the data for each measure will reveal patterns of missing data, which may indicate a problem with the administration of a particular measure. In addition, these periodic updates will allow the rate of data collection to be reviewed to determine whether the study is proceeding at a reasonable pace. The longer it takes to collect data, the greater the number of obstacles to the completion of the project. There are many projects with incomplete data sitting in file drawers.

CONCLUSIONS

When embarking on studies involving pediatric and child clinical populations, researchers must confront difficult methodological and practical issues. Careful attention should be paid to the selection of existing measures appropriate for the psychological constructs under investigation; the sample population; and monetary, logistical, and time constraints. In situations in which no measure in the

literature captures the purpose of the study or satisfies psychometric properties for the population of the current study, researchers must decide whether and how to modify existing measures and develop new ones. In general, we recommend a multitrait, multimethod approach to assessment in both research and clinical settings. Although the use of multiple measures and multiple informants is fraught with logistical and statistical difficulties, such data provide a rich source of information and can enhance the thoroughness and accuracy of psychological assessment.

REFERENCES

- Achenbach, T. M. (1978). *Research in developmental psychology*. New York: Free Press.
- Achenbach, T. M. (1991). *Integrated Guide to the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington: University of Vermont Department of Psychology.
- Achenbach, T. M. (1993). Implications of multiaxial empirically based assessment for behavior therapy with children. *Behavior Therapy*, 24, 91-116.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- Adams, D., Overholser, J.C., & Lehnert, K. (1994). Perceived family functioning and adolescent suicidal behavior. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33, 498-507.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- Beck, A. T., Schuyler, D., & Herman, I. (1974). Development of suicidal intent scales. In A.T. Beck, H. Resnick, & D. Lettieri (Eds.), *The prediction of suicide* (pp. 45-56). Bowie, MD: Charles Press.
- Bellack, A. S., & Hersen, M. (Eds.). (1984). *Research methods in clinical psychology*. New York: Pergamon.
- Bird, H. R., Gould, M. S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 78-85.
- Brinkman, D., Overholser, J., & Klier, D. (1994). Emotional distress in adolescent psychiatric inpatients: Direct and indirect measures. *Journal of Personality Assessment*, 62, 472-484.
- Brown, L., Overholser, J., Spirito, A., & Fritz, G. (1991). The correlates of planning in adolescent suicide attempts. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 95-99.
- Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Thousand Oaks: Sage.
- Crowley, S., Worchel, F., & Ash, M. (1992). Self-report, peer-report, and teacher-report measures of childhood depression: An analysis by item. *Journal of Personality Assessment*, 59, 189-203.
- Drotar, D., Agle, D. P., Eckl, C. L., & Thompson, P. A. (1996). Impact of the repressive personality style on the measurement of psychological distress in children and adolescents with chronic illness: An example from hemophilia. *Journal of Pediatric Psychology*, 21, 283-293.
- Drotar, D., Stein, R. E. K., & Perrin, E. C. (1995). Methodological issues in using the Child Behavior Checklist and its related instruments in clinical child psychology research. *Journal of Clinical Child Psychology*, 24, 184-192.
- Edelbrock, C., Costello, A. J., Dulcan, M. K., Calabro Conover, N., & Kala, R. (1986). Parent-child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry*, 27, 181-190.
- Eysenck, S. B., Easting, G., & Pearson, P. R. (1989). Age norms for impulsiveness, and empathy in children. *Personality and Individual Differences*, 5, 315-321.
- Fendrich, M., Weissman, M. M., & Warner, V. (1991). Longitudinal assessment of major depression and anxiety disorders in children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 38-42.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 12, 56-62.

- Herjanic, B., & Reich, W. (1982). Development of a structured psychiatric interview for children: Agreement between child and parent on individual symptoms. *Journal of Abnormal Child Psychology, 10*, 307-324.
- Huddleston, E. N., & Rust, J. O. (1994). A comparison of child and parent ratings of depression and anxiety in clinically referred children. *Research Communications in Psychology, Psychiatry, and Behavior, 19*, 101-112.
- Ivens, C., & Rehm, L. (1988). Assessment of childhood depression: Correspondence between reports by child, mother, and father. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 738-741.
- Jelalian, E., Spirito, A., Rasile, D., Vinnick, L., Rohrbeck, C., & Arrigan, M. (1997). Risk-taking, reported injury and perception of future injury among adolescents. *Journal of Pediatric Psychology, 22*, 513-532.
- Jensen, P. S., Traylor, J., Xenakis, S. N., & Davis, H. (1988a). Child psychopathology rating scales and interrater agreement: I. Parents' gender and psychiatric symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 442-450.
- Jensen, P. S., Xenakis, S. N., Davis, F.H., & Degroot, J. (1988b). Child psychopathology rating scales and interrater agreement: II. Child and family characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 451-461.
- Kazdin, A. E. (1980). *Research design in clinical psychology*. New York: Harper & Row.
- Kazdin, A. (1981). Assessment techniques for childhood depression: A critical appraisal. *Journal of the American Academy of Child Psychiatry, 20*, 358-375.
- Kazdin, A. E. (1992). *Methodological issues and strategies in clinical research*. Washington, DC: American Psychological Association.
- Kazdin, A., French, N., Unis, A., & Esveldt-Dawson, K. (1983). Assessment of childhood depression: Correspondence of child and parent ratings. *Journal of the American Academy of Child Psychiatry, 22*, 157-164.
- Kazdin, A., & Pettit, T. (1982). Self-report and interview measures of childhood and adolescent depression. *Journal of Child Psychology and Psychiatry, 23*, 437-457.
- Kerr, M., Hoier, T., & Versi, M. (1987). Methodological issues in childhood depression: A review of the literature. *American Journal of Orthopsychiatry, 57*, 193-198.
- Lehnert, K., Overholser, J., & Adams, D. (1996). The Cognition Rating Form: A new approach to assessing self-generated cognitions in adolescent sentence completions. *Psychological Assessment, 8*, 172-181.
- Lipovsky, J., Finch, A. J., & Belter, R. (1989). Assessment of depression in adolescents: Objective and projective measures. *Journal of Personality Assessment, 53*, 449-458.
- Loeber, R., Green, S. M., Lahey, B. B., & Stouthamer-Loeber, M. (1989). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology, 1*, 317-337.
- Loeber, R., Green, S. M., Lahey, B. B., Stouthamer-Loeber, M. (1991). Differences and similarities between children, mothers, and teachers as informants on disruptive behavior disorders. *Journal of Abnormal Child Psychology, 19*, 75-95.
- Mattison, R. E., Bagnato, S. J., & Strickler, E. (1987). Diagnostic importance of combined parent and teacher ratings on the Revised Behavior Problem Checklist. *Journal of Abnormal Child Psychology, 15*, 617-628.
- Millon, T., Green, C., & Meagher, R. (1982). *The Millon Adolescent Personality Inventory manual*. Minneapolis, MN: Interpretive Scoring Systems.
- Offer, D., Ostrov, E., & Howard, K. (1982). *The Offer Self-Image Questionnaire for adolescents: A manual*. Chicago: Michael Reese Hospital.
- Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J. E., Sanford, M., & Lipman, E. L. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*, 1078-1085.
- Overholser, J. C., Brinkman, D., Lehnert, K., & Ricciardi, A. M. (1995). Children's Depression Rating Scale—Revised: Development of a short form. *Journal of Clinical Child Psychology, 24*, 443-452.
- Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology, 20*, 51-63.
- Poznanski, E., Cook, S., & Carroll, B. (1979). A depression rating scale for children. *Pediatrics, 64*, 442-450.
- Reich, W., & Earls, F. (1987). Rules for making psychiatric diagnoses in children on the basis of multiple sources of information: Preliminary strategies. *Journal of Abnormal Child Psychology, 15*, 601-616.

- Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin, 112*, 485–499.
- Rotter, J. (1950). *Incomplete Sentences Blank: High school form [test]*. New York: Psychological Corporation.
- Silverman, E., & Overholser, J. (August 1995). *Adolescent suicidal behavior and parent perceptions of family functioning*. Presented at American Psychological Association Convention, New York City.
- Skinner, H., Steinhauer, P., & Santa-Barbara, J. (1990). *The Family Assessment Measure: Administration and interpretation guide*. University of Toronto: Applied Research Consulting House.
- Spielberger, C., Jacobs, G., Russell, S., & Crane, R. (1983). Assessment of anger: The State–Trait Anger Scale. In J. Butcher & C. Spielberger, (Eds.), *Advances in personality assessment* (Vol. 2, pp. 159–187). Hillsdale, NJ: Erlbaum.
- Spirito, A. (1996). Commentary: Pitfalls in the use of brief screening measures of coping. *Journal of Pediatric Psychology, 21*, 573–575.
- Spirito, A., Brown, L. J., Overholser, J., Fritz, G., & Bond, A. (1991a). Use of the Risk–Rescue Rating Scale with adolescent suicide attempters: A cautionary note. *Death Studies, 15*, 269–280.
- Spirito, A., Stark, L. J., Grace, N., & Stamoulis, D. (1991b). Common problems and coping strategies reported by children and adolescents. *Journal of Youth and Adolescence, 20*, 531–544.
- Spirito, A., Stark, L. J., & Williams, C. (1988). Development of a brief checklist to assess coping in pediatric patients. *Journal of Pediatric Psychology, 13*, 555–574.
- Spirito, A., Sterling, C., Donaldson, D., & Arrigan, M. (1996). Factor analysis of the Suicide Intent Scale with adolescent suicide attempters. *Journal of Personality Assessment, 67*, 90–101.
- Tarullo, L. B., Richardson, D. T., Radke-Yarrow, M., & Martinez, P. E. (1995). Multiple sources in child diagnosis: Parent–child concordance in affectively ill and well families. *Journal of Clinical Child Psychology, 24*, 173–183.
- Verhulst, F. C., & van der Ende, J. (1992). Agreement between parents' reports and adolescents' self-reports of problem behavior. *Journal of Child Psychology and Psychiatry, 33*, 1011–1023.
- Webster-Stratton, C. (1988). Mothers' and fathers' perceptions of child deviance: Roles of parent and child behaviors and parent adjustment. *Journal of Consulting and Clinical Psychology, 56*, 909–915.
- Weinberger, D. A. (1996). Distorted self-perceptions: Divergent self-reports as statistical outliers in the multimethod assessment of children's social-emotional adjustment. *Journal of Personality Assessment, 66*, 126–143.
- Weinberger, D. A., & Schwartz, G. E. (1990). Distress and restraint as superordinate dimensions of adjustment: A typological perspective. *Journal of Personality, 58*, 381–417.
- Weissman, A., & Worden, W. (1972). Risk-rescue rating in suicide assessment. *Archives of General Psychiatry, 26*, 553–560.
- White, J. L., Moffitt, T. E., Caspi, A., Bartusch, D. J., Needles, D. J., & Stouthamer-Loeber, M. (1994). Measuring impulsivity and examining its relationship to delinquency. *Journal of Abnormal Psychology, 103*, 192–205.

6

Improving Assessment in Child Clinical and Pediatric Psychology

Establishing Links to Process and Functional Outcomes

ALEXANDRA L. QUITTNER

As we approach the end of the 20th century, the field of psychology, as both a science and profession, is facing new challenges. Funding for research has been shrinking at the federal and state levels, resulting in demands for greater accountability and clearer communication with the public about how psychological research can address larger societal problems (Leviton, 1996). The national shift toward managed care is bringing similar pressures to bear on those in clinical practice who are being asked to "prove" that their treatments work (Johnstone et al., 1995; Lipsey & Wilson, 1993). Although these changes have caused significant turmoil in the research and practice communities and some real hardships (e.g., loss of income), they also may serve to redirect our energies to an area of psychology in which we have made significant and lasting contributions, namely, the development of reliable and valid measures (Applebaum, 1997). In fact, the key to addressing issues of accountability and "relevance" for both researchers and clinicians may lie in a critical examination of our approaches to assessment, with greater efforts directed toward the development of measures that elucidate important psychological processes and provide stronger links to intervention and health outcomes (O'Keefe, Quittner, & Melamed, 1996; Sechrest, McKnight, & McKnight, 1996).

ALEXANDRA L. QUITTNER • Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida 32610-0165.

An emphasis on measurement may be particularly relevant for the specialties of child clinical and pediatric psychology. This field is relatively young, and as a result has focused extensively on descriptive studies of childhood disorders and group comparisons of children with and without serious illnesses (Drotar, 1997; La Greca, 1994; Roberts, McNeal, Randall, & Roberts, 1996). Often the measures that are used are downward extensions of adult instruments (Children's Depression Inventory) (Kovacs, 1985), emphasize pathology over alterations in daily life, or contain general items (e.g., family support) that are not tied to specific functional outcomes. Researchers in the pediatric area face additional limitations because most measures have been designed for children and adolescents who are physically healthy, and the norms for these instruments are not likely to be valid for children with chronic physical conditions (La Greca, 1994; Perrin, Stein, & Drotar, 1991). Thus, there is a great need to create and refine measures that are developmentally appropriate, sensitive to the specific challenges faced by children with emotional and physical problems, and germane to the development of effective treatments (La Greca & Lemanek, 1996).

As the field moves toward intervention and outcome research, the development of theory-driven, clinically relevant measures becomes critical. Increasing the specificity and precision of measurement is likely to yield several benefits: It will facilitate the identification of targets for intervention, establish whether treatment programs lead to clinically meaningful change, and illuminate the processes or mechanisms by which change occurs (Bauman, Drotar, Leventhal, Perrin, & Pless, 1997; Weisz, Donenberg, Han, & Weiss, 1995). The purpose of this chapter is twofold: to review the limitations of current assessment approaches and suggest new directions for measurement development. First, common conceptual and methodological problems with measures in child clinical and pediatric psychology will be reviewed. This review will not be exhaustive, but rather will consider the extent to which these measures focus on process, establish links to treatment, and are useful for evaluating treatment outcomes. The second purpose of this chapter is to describe two assessment approaches that have shown promise for making contributions in these areas. Although the issues and methods discussed in this chapter have broader significance, children and adolescents with chronic illnesses and their families will be the central focus.

CONCEPTUAL LIMITATIONS OF CURRENT APPROACHES TO ASSESSMENT

Global versus Contextual Measurement

A major limitation of most measures of child and family functioning is the global level at which constructs have been defined and measured. Important risk and resilience variables, such as child temperament, stress, coping, and family functioning, have not been well-operationalized, yielding measures that are general and vague (Quittner, 1992). This has limited our understanding of the particular characteristics of the child that are problematic and the specific stressors and responses that lead to elevated risk (Drotar, 1997; Rutter, 1990). The literature on stress and coping serves as a good illustration of these problems. Although a

plethora of studies have examined the links between stress and illness in both adults and children, as well as factors that potentially mitigate these effects (e.g., coping strategies, social support), we currently have little understanding of the mechanisms underlying these effects (Coyne & Gottlieb, 1996; Quittner, Glueckauf, & Jackson, 1990). How does stress exert its negative effects? What types of stressors are most likely to place children and family members at risk? Which coping strategies are effective for specific types of problematic situations? Without answers to these process-oriented questions, it will not be possible to develop systematic interventions to address these problems.

A primary reason for our limited progress in the stress and coping area is a lack of contextual specificity in the measures that are used. Typically, studies have focused almost exclusively on global indexes of stress, such as negative life events or minor hassles (Holmes & Rahe, 1967; Kanner, Coyne, Schaefer, & Lazarus, 1981), without careful consideration of the specific demands of the stressor or its sequelae (Folkman & Lazarus, 1980). In a recent study, for example, adolescents were asked to think of the most stressful event experienced over the last month and then check off the coping strategies they used (Halstead, Johnson, & Cunningham, 1993). An adult coping measure (i.e., Ways of Coping Checklist, slightly reworded for adolescents) was then completed. The results suggested that adolescents used several coping strategies to deal with the stressful situation, indicating that most stressors are not unitary events, but rather are complex situations that include a number of different demands. This type of measure, however, does not identify which aspects of the situation were most difficult or which coping strategies were effective in managing those demands. This lack of specificity in the match between stressors and coping strategies has limited the theoretical and practical value of these studies for the purposes of intervention (Auerbach, 1989; DiGirolamo, Quittner, Ackerman, & Stevens, 1997). This decontextualization of coping research recently has been acknowledged by the originators of this assessment approach, some of whom have called for a reformulation of both the theory underlying coping research and an end to the use of coping checklists (Coyne, 1997; Coyne & Gottlieb, 1996).

Unfortunately, this approach to stress and coping measurement also has dominated the pediatric literature, yielding similarly limited information. To a great extent, general models of stress and coping (Daniels, Moos, Billings, & Miller, 1987; Lazarus & Folkman, 1984) have been utilized to assess the adaptation of children with a variety of chronic health conditions (e.g., spina bifida, diabetes, cystic fibrosis) and their parents (Kager & Holden, 1992; Mullins et al., 1991; Murch & Cohen, 1989). In many of these studies, general life stressors rather than those specific to the medical condition are assessed and tested within the stress and coping framework. In other studies, having a chronic medical condition is assumed to cause significant stress, and stress levels are not measured directly (Speechley & Noh, 1992; Varni, Wilcox, & Hanson, 1988). Finally, some studies examining parental adaptation to childhood illness have assessed general parenting stress on a standardized instrument, such as the Parenting Stress Index (Goldberg, Morris, Simmons, Fowler, & Levison, 1990; Robson, 1997). Unfortunately, these general approaches have produced conflicting results, in part, because it is not clear how stressful life events or global parenting stress are related to the immediate, daily experiences of children with serious illnesses and their parents. Chronic stressors, such as childhood illness, typically include a number of stressful tasks and de-

mands related to both the illness context (e.g., treatment regimens) and normal development (e.g., achieving independence) (Quittner et al., 1996). Measures that identify these difficulties, as well as the child and family's responses to them, are needed to increase our understanding of the processes that lead to successful adaptation.

The task of developing context-specific measures is a challenging one. A number of contextual factors must be taken into account, including the enormous developmental differences in children from birth through adolescence, the temporal course of the condition (e.g., soon after diagnosis, end-stage), and the myriad of social systems in which the child and family interact (e.g., peers, extended family, school, hospital). We have recently applied behavioral principles to the development of context-specific measures of stressors and coping strategies for children and adolescents with a chronic illness and their parents (DiGirolamo et al., 1997; Quittner et al., 1996), and these measures are now being used to guide the development and evaluation of intervention programs (Stark & Quittner, 1996; Quittner & Drotar, 1997).

Measurement of Pathology versus Functional Outcomes

A majority of studies in the child clinical and pediatric area have emphasized the assessment of "pathology" with little attention to how both psychiatric and physical disorders affect daily life. Although the diagnosis of serious emotional and behavioral disorders is fundamental to our field, an emphasis on pathology to the exclusion of other important, functional outcomes is problematic for two reasons: (1) measures of pathology do not provide problem-focused information that can be used in developing interventions, and (2) they do not illuminate the processes that lead to elevated symptomatology.

First, measures of general pathology yield only limited information about the behavioral antecedents and consequences of a clinical problem, and thus they are of limited use in guiding the development of interventions (Espelage & Quittner, 1997). For example, although an elevated score on a depression measure alerts the clinician or researcher to a problem, it does not identify the precipitants of depressive symptoms or the specific areas of functioning that have declined (e.g., academic, peer, family). Similarly, for a child who has received psychological services, what does a decrease of five points on the Children's Depression Inventory mean? Is this child now attending school regularly, completing assignments on time, and spending more time with peers? Without the assessment of concrete behaviors and activities, it is difficult to know whether the treatment has had a meaningful effect (Quittner & DiGirolamo, 1998; Sechrest et al., 1996). Even for children with psychiatric rather than physical disorders (e.g., attention-deficit hyperactivity disorder), the measurement of functional outcomes that are relevant to the performance of daily activities (e.g., completion of homework assignments, compliance with household chores), not just symptom severity, are critical for the development and evaluation of effective treatments. Thus, the assessment process should be behaviorally and functionally oriented in order to facilitate the identification of key targets for intervention. Unfortunately, global indexes of pathology do not provide this type of information.

In the pediatric area, the central focus of two decades of research has been on determining whether serious illness places children at risk for "pathological" levels of behavior problems, anxiety, or depression. Epidemiological studies in general have found that children with chronic disorders are twice as likely to have behavioral and emotional disorders as their nondisabled peers (Breslau, 1985; Cadman, Boyle, Szatmari, & Offord, 1987). A recent meta-analytic study (Bennett, 1994) also indicated that children with chronic illnesses are at slightly greater risk for depression than nondisabled peers (9% vs. 1–5%). Although these data are important for documenting the scope of the risk, as Pless and Nolan (1991) noted, "the main challenge ... has now shifted from establishing risks associated with chronic illness in general, to that of identifying specific determinants or modifiers of risk" (p. 347). These may include aspects of the larger social ecology in which the family is embedded, such as the availability of social support and community resources (Hobbs, Perrin, & Ireys, 1985).

Recent studies also have indicated that children and families who do *not* score in the clinical range on traditional measures of pathology still may be experiencing significant difficulties in managing daily interactions and medical routines (Johnson, 1995; Quittner et al., in press; Shonkoff, Hauser-Cram, Krauss, & Upshur, 1992), and could benefit from a clinical intervention (Drotar, 1997). Thus, our reliance on measures of pathology may limit our ability to identify those who are truly in need of our services.

A second problem with measures of general pathology is that they do not illuminate the processes that lead to increased risk; rather, they provide static "snapshots" of the frequency and severity of particular symptoms. For example, a parent caring for a child with a chronic illness or disability may score in the clinically elevated range on a depression measure, but this tells us nothing about why the parent is experiencing these symptoms. Is it related to the stress of providing daily care for the child; to a lack of support and communication between the parent and his or her spouse; or both?

In order to answer these questions, we need measures that are process oriented and give us information about the more proximal causes of these symptoms (Quittner et al., in press). Note, too, that questions about process shift the focus of measurement away from the individual to the larger social systems in which the person lives—immediate and extended family, the health care system, and the community. It becomes critical, then, to address not only how that parent is functioning, but how their functioning is related to their marriage, relationships with other children in the family, and connections to larger social networks (Kazak, Segal-Andrews, & Johnson, 1995; Quittner & Opiplari, 1994; Weiss, Marvin, & Pianta, 1997). What assessment procedures facilitate access to these processes?

One promising approach is the use of daily diaries. Diaries provide a window into everyday life and facilitate the measurement of activities and behaviors in "real time" (Quittner, Opiplari, Regoli, Jacobsen, & Eigen, 1992). Assessing how families organize their time each day can yield important and highly accurate estimates of behaviors, such as the amount of parental time and attention directed to siblings or the frequency with which medical routines are performed (Freund, Johnson, Silverstein, & Thomas, 1991; Quittner & Opiplari, 1994). Thus, they may be used to assess both complex family processes and changes that occur in daily activities and behaviors as a result of intervention (Quittner et al., in press; Reed,

Richards, Moneta, Holmbeck & Duckett, 1996; Shiffman, Paty, Gnys, Kassel, & Hickcox, 1996). Although activity pattern assessment and daily diaries have played an important role in the rehabilitation field and in research on coercive family processes (Chamberlain & Reid, 1987; Rock, Fordyne, Brockway, Bergman, & Spengler, 1984), they have been underutilized in child clinical and pediatric research (see Csikszentmihalyi & Larson, 1984, Montemayor, 1982 for exceptions).

METHODOLOGICAL LIMITATIONS OF CURRENT APPROACHES TO ASSESSMENT

Using Measures Standardized with Healthy Samples

Most standardized measures of psychopathology were developed and normed with samples of nondisabled children and children being referred for mental health problems. Thus, the appropriateness of their use for children with chronic illnesses is uncertain (La Greca, 1994; Perrin et al., 1991). Although using a normed instrument typically strengthens the results of an evaluation, several problems have been identified with the use of these measures for children with chronic physical conditions.

First, it is difficult to know how to interpret elevated scores that include items assessing physical symptoms. For example, the most widely used behavior checklist for children and adolescents (Child Behavior Checklist) (Achenbach, 1991) contains items that refer specifically to a medical diagnosis (e.g., asthma, allergies) as well as a variety of physical symptoms (e.g., "feels dizzy," "lacks energy"). For children with chronic illnesses or disabilities, these items reflect legitimate aspects of their medical condition rather than evidence of disturbed behavior. If parents and teachers endorse these items, then that child's behavior problem scores are likely to be elevated and an erroneous conclusion may be drawn that his or her behavior problems are above the "norm." Although it may be easier to recognize this error in a clinical setting, in a research context, these elevated scores may be interpreted to mean that children with chronic illnesses have higher rates of behavior problems than those who are nondisabled (Perrin et al., 1991).

Similar problems may be noted for other standardized measures, such as the Children's Depression Inventory (Kovacs, 1985). When completed by a child with a chronic physical condition, several items may reflect manifestations of the illness rather than symptoms of pathology (e.g., feeling tired; not feeling like eating; worrying about aches and pains). Inclusion of these items again may lead to spuriously high depression scores for this population. Thus, information provided by these standardized measures may be only marginally useful for intervention and may provide inaccurate data on the effectiveness of treatment.

The immediate solutions to these problems are not obvious. If items reflecting physical symptomatology are deleted when the measure is scored, then the normative data and clinical cutoff scores cannot be used. This eliminates one clear advantage of the standardized assessment approach. In addition, if the measure is being used to compare levels of depression in children with and without chronic illnesses, then removing these items poses another set of problems for analysis and interpretation of the results. As suggested throughout this chapter, the assessment

of "pathology" per se has conceptual as well as methodological limitations. At a minimum, we need to be cautious about using these measures and supplement them with instruments that do not contain confounded items.

Sensitivity and Specificity of Measurement

A second problem with the use of standardized measures for children with chronic medical conditions is their lack of sensitivity to behavioral variations within the normal range. These measures were designed to identify children and adolescents with serious pathology, and as was mentioned earlier, the vast majority of children with chronic illnesses do not evidence "pathology" as it has been traditionally defined. However, they are more likely than nondisabled children to experience disruptions in daily functioning, such as school attendance, academic performance, and peer relationships (DiGirolamo et al., 1997; Lavigne & Faier-Routman, 1992). These challenges may increase symptoms of anxiety, depression, and social isolation, but not to the extent that the child's score falls within the clinical range. Thus, the use of measures that are relatively insensitive to these minor but important fluctuations in symptomatology may lead to both the loss of valuable information and the denial of psychological services to children who are at risk (Drotar, 1997).

A reliance on standardized measures of pathology leads to a related problem: a lack of specificity about the problems that are most relevant for children with chronic physical conditions. A majority of assessment tools used in the child clinical and pediatric areas assess global constructs, such as self-esteem, behavioral adjustment, and psychological distress. However, given that children with chronic illnesses experience difficulties that are closely linked to their medical condition (e.g., taking medication, missing days at school, explaining their illness), the measures we use must assess these more mundane but important issues of daily living (DiGirolamo et al., 1997; Quittner et al., 1996).

Similar limitations may be noted for self-report measures of family functioning (e.g., Family Environment Scale) (Moos & Moos, 1981). These questionnaires typically yield general indexes of cohesion and conflict, but do not assess the issues that are critical for the family, such as conflicts about adherence to daily medical regimens or adolescent efforts to become independent (DiGirolamo et al., 1997; Wysocki, 1993; Wysocki, White, Bubb, Harris, & Greco, 1995). The lack of relevant, contextually based measurement has led to conflicting findings and few direct links between assessment and treatment. It has been the primary stumbling block in our efforts to move beyond description to a deeper understanding of the processes that lead to successful or unsuccessful family adaptation.

Units of Measurement and Analysis

A final problem that has plagued studies of child clinical and pediatric populations is confusion about the units of measurement and analysis. A majority of studies investigating the impact of childhood disorder on the family have used an *individual's* report to assess various *family-level* constructs (e.g., cohesion, conflict). In a recent review of health-related studies in family journals, Patterson (1990) found that 80% of the studies from 1980 to 1989 relied on a single family

member's account of the family group; in most cases, the measure was some type of self-report questionnaire. This blurring of the distinction between individual and family levels of assessment is problematic because of potential differences in how individuals within the family view the situation (Glueckauf, *in press*; Quittner et al., *in press*). For example, the same stressful context is likely to affect family members in different ways; thus eliciting a "family-level" report of conflict or cohesion by asking one member of the family may not be valid. As Patterson (1990) notes, "asking the individual how the family copes appears really to measure how the individual copes" (p. 416).

What are the alternatives? One possibility is to calculate a relational score by combining the individual reports of several family members (Patterson, 1990). A family-level score of cohesion could be created, by adding or multiplying together various family members' reports on this dimension of family functioning. However, it is not clear that summing reports from individual family members is the same as an assessment of the family unit (Lewis, Beavers, Gossett, & Phillips, 1976; Reiss, 1981).

A better option is to directly observe interactions among members of the family unit or develop assessment tools that measure family-level functioning (Glueckauf et al., 1992; Hauser et al., 1993). The programmatic research of Hauser and colleagues (Hauser, 1990; Hauser et al., 1993) serves as an excellent example of observational assessment with families caring for an adolescent with diabetes. Families engage in a semistructured interview (*Family Life Events Interview*, adapted from Reiss and Oliveri, 1980) about how a recent major event (e.g., diagnosis of the adolescent's diabetes) has interrupted or changed their activities and interactions as a family. The family interview is audiotaped, transcribed, and then coded by independent raters for the occurrence of various appraisal and coping processes. Similar "family discussions" are being used as outcome measures of family interventions for adolescents with diabetes (Wysocki et al., 1997), cystic fibrosis (Quittner & Drotar, 1997), and attention-deficit disorder (Robin, 1998).

Diary methods also hold promise for quantifying family-level interactions and processes without the potential limitations of direct observation (e.g., social desirability responding). Diary measures, particularly those that are completed on-line (e.g., via telephone, palm-top computer) rather than by mail, have yielded valuable information about the frequency and intensity of conflict in the home (Chamberlain & Reid, 1987; Reed et al., 1996), the negotiation of responsibilities for medical treatment (Freund et al., 1991; Quittner et al., 1992), and amounts of time spent in recreational and child-focused play (Quittner et al., *in press*).

Summary

In sum, several key problems have been identified with current assessment approaches. In particular, there has been an emphasis on the measurement of global constructs that are not linked to the specific clinical, developmental, and family contexts in which the child or adolescent lives. This has severely limited our ability to understand the processes underlying adaptation to childhood disorder and disability and to develop interventions that are likely to produce meaningful change. Increasing the contextual specificity of the measures we use does involve trade-offs, including restrictions on the generalizability of our findings across

developmental ages and clinical populations (Drotar, 1994). However, the shift toward treatment outcome research requires this level of specificity in order to develop closer links between assessment, treatment, and outcome.

A CONTEXTUAL APPROACH TO MEASURING STRESS AND COPING PROCESSES

Research on stress and coping has been one of the most active areas of research for psychologists over the past 15 years. Literally hundreds of studies have assessed the links between stressful events, the use of various coping styles or efforts, and physical or psychological outcomes. In most stress models, coping processes are a central focus because they mediate the relationship between the stress and adaptation (Lazarus & Folkman, 1984; Quittner, et al., 1990) and serve as potentially important targets for intervention. However, despite considerable interest and effort, we currently have little understanding of how and under what conditions coping strategies affect physical and psychological outcomes, and even less empirical evidence to suggest that they are effective!

As mentioned earlier, the major reason for our limited progress is a lack of contextual specificity in the measurement of both the stressor and coping strategy. The predominant methodology for studying these processes, in both adults and adolescents, has involved completion of a list of stressful life events or hassles, or identification of a stressful situation that has occurred during a prescribed interval (e.g., 1 month to 1 year). Next, a "coping checklist" is completed that includes a number of predetermined strategies ("I am inspired to do something creative"). Thus, across subjects there is no correspondence between the stressful situations they recall and the coping strategies that are endorsed. This has made it virtually impossible to draw any conclusions about which strategies are effective for particular situations (Coyne & Gottlieb, 1996).

Additional problems with the measurement of stress have been noted. First, subjects are asked to report on stressful situations that have occurred over the past year, or even month, and then remember what they were thinking and feeling. This retrospective type of assessment has led to serious problems with reliability and recall (Ptacek, Smith, Espe, & Raffety, 1994; Raphael, Cloitre, & Dohrenwend, 1991). Second, respondents completing these checklists may endorse very different types of events, yet receive a similar "stress score" (Hobfall, 1989). Finally, these checklists provide minimal information about the contextual aspects of the event itself, such as when it occurred, who was involved, and what demands it placed on the individual (Menaghan, 1983; Pearlin & Turner, 1987).

In particular, little attention has been given to the chronicity of the stressor, despite evidence that ongoing stressors, as opposed to discrete events, are more strongly related to negative outcomes, such as depression and family disruptions (Quittner et al., 1990; Sandler, Wolchik, & Braver, 1988; Wheaton, 1997). Chronic stressors, such as childhood illness, typically include a number of stressful tasks and demands related to both the illness (e.g., treatment regimens) and to normal development (e.g., completing schoolwork), and thus present a complex set of ongoing challenges for the family that cannot be reduced to a single "event." In addition, chronic stressors may elicit different patterns of coping behavior, such as

anticipatory efforts designed to avert a more serious crisis (e.g., careful management of medications) or habitual, routinized strategies that have been refined over time in response to repeated demands (Coyne & Gottlieb, 1996). Clearly, new methodologies and assessment tools are needed to capture these more complex processes and outcomes (DiGirolamo et al., 1997; Gignac & Gottlieb, 1996).

Similar problems have plagued the measurement of coping. Respondents are presented with predetermined lists of strategies that are global in nature and may or may not be relevant to the stressful situation (e.g., "find new faith"). This lack of specificity in the match between the stressor and coping strategy has led to inconsistent findings regarding which types of strategies are effective, and few practical implications for intervention. Further, several studies have indicated that children and adolescents use different coping strategies as a function of their developmental age, since age directly influences the cognitive, emotional, and social skills that are available (Band & Weisz, 1990; Spirito, Stark, & Tyc, 1994). Thus, the use of generic coping checklists may obscure important developmental and contextual shifts in coping behavior.

Finally, a major thrust of coping research has been the development of a typology of coping strategies, typically divided into two broad categories: problem-focused and emotion-focused coping (Folkman & Lazarus, 1980). Two problems with this approach may be noted. First, coping strategies are categorized into these two domains prior to their use in a specific situation. This raises an important question: Can the function of a coping strategy be determined outside of its context? For example, although strategies that include distraction (e.g., relaxation techniques) are usually classified as emotion-focused, they also can be taught as useful strategies to reduce stress, which places them in the problem-focused category (Auerbach, 1989). Similarly, Meyerowitz, Heinrich, and Schag (1983) concluded that denial, also categorized as emotion-focused, was often effective for adjusting to some aspects of having cancer. Further, although the original transactional theory of coping hypothesized that adaptation to stress was an ongoing process, little attention has been given to these temporal relations. For example, early in the stress process, strategies such as denial may be highly effective because they allow an individual time to think, plan, and process complex information. The continued use of this strategy, however, may lead to withdrawal and social isolation (Coyne & Gottlieb, 1996).

A second problem with this typology is the assumption that a rational attempt to master and "fix" a problem leads to positive adaptation, whereas the use of emotion-focused strategies leads to greater distress. The empirical basis for this claim is not strong. In both the adult and child literatures, inconsistent relationships have been found between the use of problem-focused coping and decreased psychological distress (Compas, Malcarne, & Fondacaro, 1988; Revenson & Felton, 1989; Wertlieb, Weigel, & Feldstein, 1987). Could the value placed on problem-solving strategies be related, in part, to our cultural values? And is the overwhelming tendency of respondents to endorse action-focused strategies and rate them as effective influenced by social desirability responding (Coyne, Sonnega, & Carter, 1993; Vaux, 1985)? These problems have led researchers to question the goal of identifying generally effective coping strategies, given that the effectiveness of a coping response cannot be determined *independently of the context in which it is*

used (Quittner et al., 1996; Spirito, 1996; Stone, Greenberg, Kennedy-Moore, & Newman, 1991).

Application of the Behavior Analytic Model to the Assessment of Stressors and Coping Strategies

In an effort to reconceptualize the coping process and its measurement, we recently conducted a series of studies applying the behavior-analytic model advocated by Goldfried and D'Zurilla (1969) to the assessment and evaluation of coping strategies in children and adolescents with a chronic illness and their parents (DiGirolamo et al., 1997; Quittner et al., 1996). We chose the behavior-analytic model because it embeds the assessment process within the context in which an individual interacts and attempts to solve problems. In light of the problems with the coping literature cited above, it also offers several methodological advantages for developing rigorous measures of coping: (1) stressors are conceptualized and measured as problems occurring in specific situations; (2) coping strategies are assessed as the responses (e.g., cognitive, behavioral, emotional) to those problem situations; (3) the effectiveness of a particular coping response is determined within that context, as judged by those most familiar with the problem; and (4) the interactive nature of the coping process can be captured in the problematic vignettes that are created.

In these studies, we have eschewed the arbitrary categorization of coping strategies into particular classes (e.g., problem- and emotion-focused). Instead, our premise is that thoughts and behaviors can be judged as competent or incompetent only by examining them within the context of the frequent and difficult situations presented to individuals in their environments. Within this framework, situations are considered problematic (or stressful) "if no effective response alternative is immediately available to the individual confronted with the situation" (D'Zurilla & Goldfried, 1971, p. 108). Thus, a problematic situation challenges the individual to generate a response that is most effective for that situation. Competent behavior (effective coping) is defined as a "response, or pattern of responses, that alters the situation so that it is no longer problematic, and ... produces a maximum of other positive consequences or a minimum of negative ones" (Goldfried & D'Zurilla, 1969, p. 168). In contrast to the usual coping typology, this model allows us to test empirically whether a specific strategy "works"; if it is effective, it should reduce the frequency or difficulty of that problematic situation or minimize future negative consequences (e.g., hospitalizations).

This model begins with an in-depth analysis of the situations that are most common and difficult for a specific population, and is followed by an assessment of the responses commonly made to these situations and then judgments of the effectiveness or competence of these responses. A unique advantage of this assessment model is the inclusion of the "insider's perspective" in all phases of the measurement process. This is somewhat analogous to obtaining input from the consumer when developing a new treatment program or set of services (Glueckauf, 1990). It is very helpful to know, from the perspective of the target population, what services are needed and how to best deliver them. In the context of measurement development, we wanted to know from the "experts," that is, the children and

parents who cope with a particular chronic illness each day, which problematic situations were most relevant, and which coping strategies were most likely to succeed. This is likely to increase both the clinical relevance and validity of the final measure.

Five phases of measurement development are included: situational analysis, item development, response enumeration, response evaluation, and instrument development. Each phase is described briefly below. The final measure consists of a role-play inventory of the most common and difficult problems, as well as empirically derived criteria for evaluating the competence of specific coping strategies (Quittner, 1996).

Situational Analysis

The primary goal of this phase is to collect a large representative sample of concrete problematic situations that are likely to confront a specific population. The focus is on situations that involve specific, meaningful, and frequently occurring problems that require an effective coping response. Detailed descriptions of problem situations are elicited, including important contextual details (e.g., setting, time of day), and ratings of the frequency and difficulty of these problems. Data are collected from individuals in the environment who are most familiar with the problems confronted by this population (e.g., adolescents who have the illness, health care professionals who provide care) and who are most likely to provide consequences for competent and incompetent behavior (e.g., parents). Multiple methods of data collection (e.g., structured interviews, daily diaries) have been recommended to produce greater contextual detail and to elicit problems that are ongoing in nature (Goldfried & D'Zurilla, 1969; Turk, Sobel, Follick, & Youkilis, 1980). Because the majority of problematic situations encountered by our populations (i.e., children, adolescents, parents) are interactional—either dyadic or triadic (e.g., parents reminding teen to do treatment; friends teasing the child)—we expanded the behavior-analytic model to include sequences of responses (e.g., “What happened next?”). Common sequences were then incorporated into the role-play vignettes.

Item Development

Next, the problematic situations are transcribed and content-analyzed. Similar situations are grouped together and statistical analyses are used to determine which situations occur most frequently and are rated most difficult across respondents. Situations that are redundant, infrequent, or not very difficult are eliminated. Decision rules, based on the distribution of these scores, are then applied to select the most relevant (i.e., salient) problematic situations. These situations are then rewritten in the form of role-play vignettes.

Response Enumeration

The purpose of this phase is to elicit a wide range of coping responses (i.e., from very effective to very ineffective) to the audiotaped vignettes generated earlier.

Participants are presented with each problematic situation and asked to place themselves in that situation and describe how they would respond.

Response Evaluation

In this phase, the effectiveness of each coping response is determined. A panel of "expert judges" are recruited who are familiar with the problems facing this population (e.g., children with the illness, pediatric psychologists) and they rate the effectiveness of each strategy on a scale from extremely competent to extremely incompetent. In making their evaluation, judges are asked to consider (1) how effective the response is in reducing the frequency or difficulty of the problem, and (2) how likely it is to decrease the probability of future negative consequences. Judges also provide a rationale for their ratings and a description of the elements that make up a highly competent response.

Instrument Development

The purpose of this final phase is to refine the measure (e.g., delete vignettes that do not elicit a range of responses) and develop a manual that can be used to rate participants' coping responses. The final manual consists of written descriptions of each vignette, scoring criteria for judging response efficacy, and examples of each response type (i.e., extremely competent to extremely incompetent) (Quittner, 1996).

Role-Play Inventory of Situations and Coping Strategies

Over the past 6 years, we have been applying this measurement process to two developmental cohorts of children with a serious, chronic illness and their parents: adolescents ages 11 to 17 with cystic fibrosis (CF), parents of adolescents with CF, school-age children 6 to 11 with CF, and parents of children with CF. To date, we have completed all five phases of measurement development with three of the four samples: adolescents, parents of adolescents, and school-age children (DiGirolamo et al., 1997; Quittner et al., 1996; Wager & Quittner, 1997). To illustrate how these measures were developed and how they may be used in an intervention program, a brief summary of how we developed the Role-Play Inventory of Situations and Coping Strategies (RISCS) measure for parents of adolescents with CF appears below (for information about the RISCS for teens and parents of school-age children see DiGirolamo et al., 1997; Quittner et al., 1996).

To develop a comprehensive inventory of the problematic situations encountered by parents of adolescents with CF, we conducted home interviews or daily phone diaries (Quittner & Espelage, 1997) with 84 parents (45 mothers and 39 fathers) at two major medical centers (i.e., Indianapolis and Pittsburgh). In addition, seven health care professionals (e.g., nurses and pulmonologists working on a CF team) completed a structured interview. Both CF-specific and normal developmental issues were assessed and respondents were asked to provide detailed information about the context (setting), the individuals involved (e.g., teen and sibling), and any interactional sequences that occurred. We also obtained ratings of the frequency and difficulty of each problematic situation. A total of 1638 problem-

atic situations were elicited from parents and an additional 67 situations were obtained from health care professionals.

Next, these problematic situations were content-analyzed into 11 domains (e.g., discipline, medications, and treatment). Using this category scheme, two independent raters then recategorized the problem situations, with interrater agreement above 90%. The frequency and difficulty ratings were then used to establish criteria for selecting the most relevant situations. Twenty-eight problematic situations met these criteria and were subsequently written in the form of role-play vignettes, with slightly different wording for sons and daughters. The following is a sample vignette from the medications and treatment domains:

You know it is time to do treatment, and when you start getting the mist ready, your son says, "Gee, I'm tired, can't we skip it just this one time?" He always wants to get out of it, and you find it frustrating. What would you say or do in this situation?

Problematic situations related to typical parent-teen conflict also emerged:

It is Friday night and you agreed to let your daughter go out with some friends, as long as she followed her regular curfew. She is now over an hour late, and you are feeling worried and angry. Finally, she walks in the door. What would you say or do in this situation?

During the response enumeration phase, the 28 audiotaped vignettes were then presented to parents at both CF centers to elicit their coping responses. A total of 2352 different responses were tape-recorded and transcribed for presentation to the expert judges (response evaluation phase). Twenty-six judges (i.e., 13 parents of adolescents with CF and 13 health care professionals familiar with the disease) were recruited and trained to evaluate the competence of the coping strategies. Based on the judges' ratings, the role-play measure was refined and their ratings were used to establish criteria for scoring the competence of responses to each vignette (instrument development). As expected, we found preliminary evidence of convergence between our measure of coping efficacy and other measures, particularly when the context was "matched" (e.g., parents who used more effective strategies for discipline problems reported that their teens were more socially skilled and had few externalizing behavior problems) (Wager & Quittner, 1997). More evidence is needed, however, that competent responses to vignettes on the RISCS converge with effective coping behaviors observed by others (e.g., parents, teachers) in the environment.

How can the RISCS measure be used? First, it provides detailed information about the problems that are most frequent and difficult for a specific population. Having this information is critical to developing an intervention that addresses the most relevant problems and outcomes. The RISCS measure also identifies strategies that are effective for coping with these problems. Rather than relying on the usual "typology" of problem- and emotion-focused coping, this measure outlines *empirical* criteria for judging the efficacy of a coping response derived from those most intimately familiar with the scope of the problems, the range of possible responses, and the long-term consequences of utilizing particular strategies. Incorporation of the consumer's perspective increases the validity of the assessment process and the credibility of the researcher or clinician who uses it (Glueckauf, 1993). Imagine discussing dating problems with a teen and being able to say that, according to a

large number of teens like you, this is what works best! We are currently using the information generated from the RISCS measures to design a family intervention program for adolescents with CF and their parents (Quittner & Drotar, 1997). These measures will serve as critical outcome variables.

Another advantage of this assessment system is that it provides information on the specific strengths and weaknesses of individuals. Interestingly, our data suggest that parents and teens are able to cope effectively with some situations and not others (Quittner & DiGirolamo, 1998). This is consistent with the premise that coping responses are likely to vary with the demands of the situation and the nature of the interaction. Thus, the RISCS measure holds promise for generating individual profiles of competence for our samples of children, adolescents, and parents.

Figure 1 illustrates the competence of a mother's and father's coping strategies across domains; their daughter is an adolescent with CF. As can be seen, the mother employs highly competent strategies for problems in the treatment, medical care, and discipline domains, but is coping less effectively with problems related to dietary and health issues. In contrast, the father is not coping effectively with problems in the treatment or discipline domains, but, like his wife, is responding more competently to problems related to medical care. Using data from this assessment system, an intervention can be tailored to the specific needs and strengths of family members. Evidence from other intervention studies indicates that targeted interventions are more likely to effect meaningful change than more globally applied programs (Weisz et al., 1995). In sum, this measure holds promise for linking the assessment and treatment process and for quantifying, in specific and functional terms, the outcomes of treatment.

DIARY METHODS OF ASSESSMENT

Diary methods of data collection also have shown considerable promise for advancing our knowledge of family processes and documenting treatment outcomes. In addition, they have the potential to address many of the limitations of current measurement approaches, including their lack of specificity, inadequate attention to process issues, and weak links to functional outcomes. Diary methods are increasingly being used to assess a wide range of important child and family behaviors (e.g., parenting practices, activity patterns), and offer several advantages over traditional self-report techniques. First, daily diaries allow us to measure, in a reliable and nonreactive manner, an ongoing series of behaviors and interactions that would not be accessible through either questionnaires or in vivo observations (e.g., coercive family interactions, adherence to medical treatment) (Chamberlain & Reid, 1987; Quittner et al., in press; Reynolds, Johnson, & Silverstein, 1990).

Second, diary assessments that are conducted on-line as target behaviors are occurring (e.g., urges to smoke) or soon after (e.g., nightly phone calls) are far less susceptible to memory and recall problems, as well as confabulated or missing data, than paper-pencil diaries collected after some specified interval (e.g., 1 week) (see Stone & Shiffman, 1994, for a review of these issues). These more rigorous diary methods typically utilize technologies, such as telephones, beepers, and palm-top computers, that have been shown to facilitate recall and minimize the biases associated with autobiographical memory (Stone, Kessler, & Haythornthwaite, 1991).

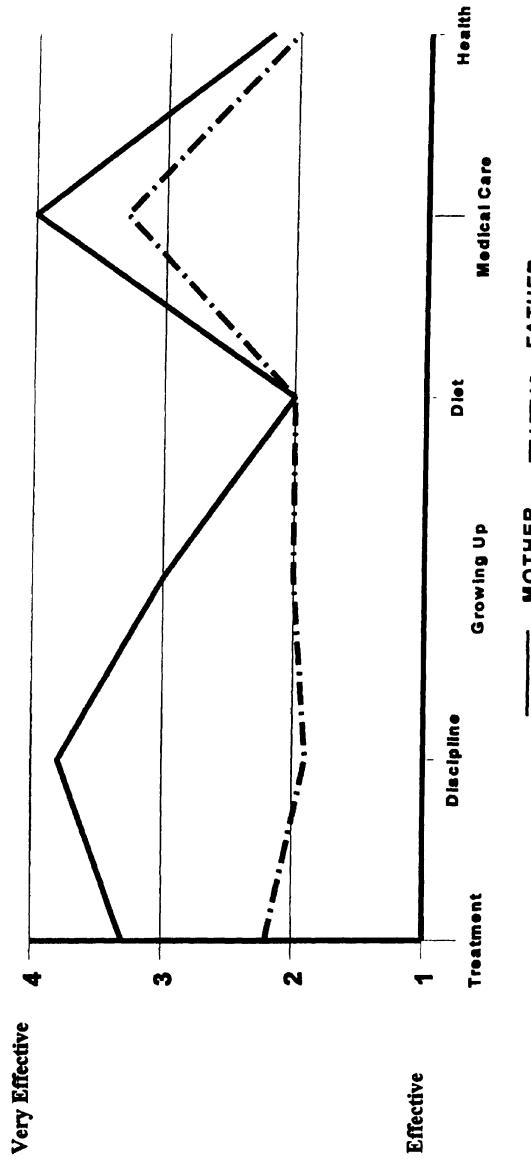


Figure 1. Profiles of coping efficacy derived from the Role-Play Inventory of Situations and Coping Strategies measure for a mother and father who have a female adolescent with cystic fibrosis.

Diary methods also offer the investigator a tool for measuring potentially sensitive behavior patterns (e.g., differential treatment of siblings) in an unobtrusive way (e.g., Quittner & Opiari, 1994). Tracking how time is spent in various activities throughout the day can yield important and accurate estimates of behaviors, such as playtime with a sibling or compliance with medical routines, without respondents being aware of the behaviors of interest. This can substantially decrease biases, such as social desirability responding (Quittner et al., in press).

Finally, diary procedures offer the investigator tremendous flexibility in the types of variables that can be measured (e.g., conflict during meals, a couple's division of household labor), and they can be easily modified to change the focus of assessment. For example, using our daily phone diary procedure, we can obtain data on who is performing medical treatments on a given day (e.g., mother or father) or on the amount of time spent on treatments. In particular, diary methods lend themselves to the assessment of family-level processes, since most activities involve several members of the family (e.g., recreation with a spouse, parent-teen arguments). This "window" into daily family life can help us understand how families negotiate their roles and responsibilities and may provide critical information about changes in behaviors and activities that occur as a result of an intervention (Quittner et al., in press).

Daily Phone Diary Procedure (DPD)

We have recently developed a Daily Phone Diary (DPD) procedure that may be used to track parents and children through their activities, interactions, and mood. This process entails a phone call in the evening to complete the DPD. Using a cued-recall procedure, respondents report on all activities lasting 5 minutes or longer over the past 24 hours. The following information is obtained: the type of activity, duration, purpose (i.e., instrumental or recreational), who was present, and mood (i.e., positive to negative). This procedure has yielded high levels of interrater agreement for the categorization of activities (above 90%), considerable stability in reports of activities over a 3- to 8-week period, and reliable differences in activity patterns for week versus weekend days (Quittner & Espelage, 1997; Quittner & Opiari, 1994). Contrary to popular belief, the diary procedure is not overly time-consuming (i.e., about 20 minutes per phone call) and produces a richness of description that is lacking in many other assessment methods.

To date, we have used the DPD to measure a variety of behaviors and family interactions, and we are now using it as an outcome measure to assess changes in treatment adherence and family conflict that have been targeted in an intervention (Quittner & Drotar, 1997). To date, the DPD procedure has been used to: (1) assess the extent to which parents provide time and attention differentially to siblings; (2) assess the frequency and difficulty of daily problems for school-age and adolescent children with CF; (3) compare patterns of child care and recreation in families with and without a chronically ill child; and (4) assess role strain and the exchange of affection in couples caring for children with and without a chronic illness (Quittner & Opiari, 1994; Quittner et al., 1992, in press).

In our experience, the diary measure provides a window into the complex alterations that may occur in families who have a child with a serious illness or disability. Rather than finding that all areas of individual and family functioning are affected negatively by a child's illness (the deficit or pathology model), our diary

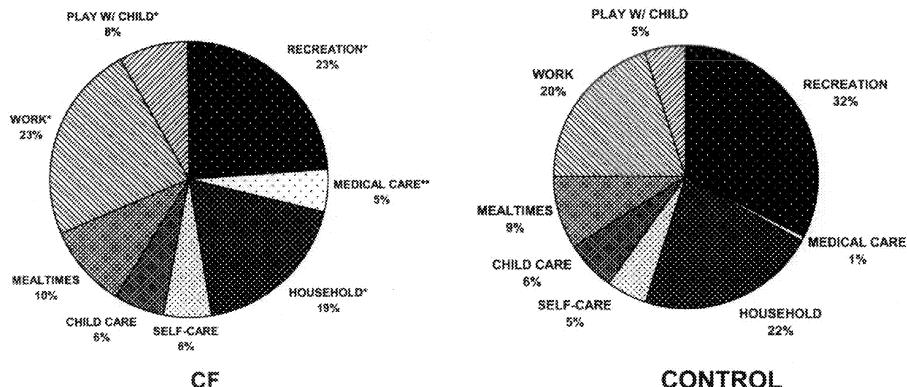


Figure 2. Activity patterns derived from the Daily Phone Diary with 33 couples caring for a child with cystic fibrosis and 33 couples caring for a nondisabled child.

data suggest that differences in daily life are specific and closely linked to the child's developmental and illness-specific needs. For example, Fig. 2 compares the activity patterns of couples caring for a young child (ages 2 to 6) with and without CF. Surprisingly, the two groups of families spend similar amounts of time in child care, self-care, and mealtime activities. Couples who have a child with CF spend more time performing daily medical treatments and less time in recreation; however, they spend more time than comparison couples in child-focused play! Thus, families appear to make complex trade-offs in order to accommodate the demands of a chronic illness. Given that we also found a positive relationship between time for recreation and psychological well-being and marital satisfaction, the clinical implications of these data are clear: couples with substantial caregiving responsibilities may need increased time for recreation (e.g., weekend relief, trained babysitters) and assistance in achieving a more equitable division of treatment responsibilities (Quittner et al., in press).

In sum, the DPD provides information about the processes that lead to role strains in families experiencing chronic stress and yields data that can be used to develop potentially effective interventions. Recently, we have developed a computerized version of the DPD that automates the tracking procedure and inputs the data directly into a computer file. This ensures that a full 24 hours of data are collected and eliminates the possibility of incomplete or missing data. The DPD is now being used by other investigators to assess how teens and their families are recovering from traumatic brain injury and to identify triggers for episodes of recurrent abdominal pain.

CONCLUSIONS

In this chapter, the limitations of current measures and assessment approaches in child clinical and pediatric psychology have been reviewed. Central to these problems is the overuse of global constructs that have not been clearly operationalized and measured and a lack of attention to the processes that link important

psychological variables to health outcomes. There is a need to develop measures that focus on important activities of daily living and on functional outcomes. How can we move forward and address these problems in terms of both graduate training and future research?

Implications for Training

In order to revitalize the assessment area and stimulate new research, we must do a better job of training our graduate students in the principles of psychometric theory and the appropriate use of assessment tools. There is evidence that quantitative training in doctoral programs has decreased generally (Aiken, West, Sechrest, & Reno, 1990), and that this is particularly true with respect to training in measurement. Although most programs require students to take at least one assessment course, the emphasis is often on “testing” as opposed to psychometric theory (Sechrest, 1993). The fundamentals of reliability and validity, including various subtypes, tend to be glossed over and graduate students spend little time actually calculating these coefficients. Without hands-on experience in applying these principles and seeing their impact on assessment data, it may be difficult to convince graduate students of their critical importance in the research process.

We also need to train graduate students in new theoretical developments and techniques in the measurement area. More recent theories, such as item response theory and G theory, and the emergence of creative measurement techniques (e.g., multidimensional scaling, experience sampling methods) have the potential to contribute greatly to our understanding of assessment data. Students should be exposed to these topics in graduate school and encouraged to use them in their research (Tinsley, 1992). At a minimum, they should be conversant with these theories so they can critically evaluate their use in published research. Graduate programs tend to reflect the knowledge base of their faculty, and if those faculty do not keep up with new developments, then progress and innovation in the field may be slowed.

As a faculty member in a research-oriented clinical program, another problem I have observed is that graduate students tend to rely on “off the shelf” measures and often do not think critically or creatively about measurement issues when designing their own research. In the child clinical and pediatric area, this has resulted in the overuse of self-report measures and the proliferation of literally hundreds of different questionnaires! In training graduate students, we need to emphasize the direct link between the quality of the measures we use and the integrity and meaningfulness of the data we collect. That may translate into greater use of behavioral and functional measures, which may increase the demands on a student’s time or energy. However, there is nothing more frustrating for a student (or an advisor) than spending a significant amount of time and effort on data collection that yields confusing or uninterpretable results. If graduate programs require smaller research projects leading up to the dissertation, these might serve as useful opportunities for piloting and evaluating new measures.

Future Research Directions

In this chapter, I have advocated the development of measures that are more contextual, that is, embedded in the specific situations and interactions that are

relevant to the population. Taking a more contextual approach is admittedly complex, since many factors must be considered, such as the child's age, the type of psychological or physical disorder, the onset and course of the condition, and the various social systems in which the child and family interact. The question that is inevitably asked is: How can we develop situation-specific measures for each of the major chronic conditions? And if we develop such specialized measures, how can we generalize our findings across conditions?*

This is a difficult question. On the one hand, because of limited time and resources, it might seem wise to develop generic measures that are applicable to a variety of chronic childhood conditions. However, my reviews of the literature suggest that attempts to do this have produced vague and inadequate answers to such basic questions as: When are children and families likely to experience adjustment problems? When and under what conditions do children adhere to their treatment regimen? The clinical utility of these studies is also not clear, since for the most part they do not translate easily into treatment recommendations or intervention programs (Quittner & DiGirolamo, 1998).

On the other hand, if our goal is to develop comprehensive, replicable data on child and family adaptation to chronic conditions, with clear implications for treatment, then perhaps the investment in studies that are more contextual and process oriented will yield clearer payoffs. This is an empirical question and the next step in our program of research. As we begin to use the RISCS measure to identify important, context-specific goals for the treatment of adolescents with CF and their parents, we will be able to document our success in addressing and ameliorating the specific needs of these families. Can we effect statistically and clinically significant behavioral change? If so, is it reflected in our measures? And, as we track changes in the daily lives of these families through the diary approach, can we identify the processes underlying these changes? In balancing the trade-offs of generalizability and specificity, it may be that taking an inductive rather than deductive approach is what is needed. Once we have a solid knowledge base on specific conditions, we may then be able to identify patterns across conditions.

In sum, as we move into the next century, psychology is uniquely positioned to contribute to the advancement of science and practice in both the mental health and medical fields. Our expertise in measurement development, psychometric theory, and evaluation is equaled by none. Greater support and encouragement should be given to creative efforts in the measurement domain, and we should continue to promote graduate education and training relevant to this area. Psychologists are also at the forefront in terms of the utilization of new technologies for assessment and treatment. These technologies will become increasingly important as we focus on functional assessment and service delivery (Glueckauf, 1996). Thus, we are in a strong position to move forward and address the current and upcoming challenges of relevance, efficacy, and accountability.

ACKNOWLEDGMENTS. Preparation of this chapter was supported in part by grants from the National Institutes of Health (FIRST Award R29 HL47064 and K04

*Note that health care professionals in other areas (e.g., pediatricians, occupational therapists) always tailor their treatment recommendations to the specific illness or clinical problem, and would not treat a child with asthma the same way they would treat a child with cystic fibrosis.

HL02892). Special thanks to Laura Wager, Robert L. Glueckauf, and Dorothy Espelage for their comments on an earlier draft of this manuscript.

REFERENCES

- Applebaum, M. I. (1997, April). *Quantitative perspectives on development*. Symposium conducted at the biennial meeting of the Society for Research in Child Development, Washington, DC.
- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Aiken, L., West, S. G., Sechrest, L., & Reno, R. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of PhD programs in North America. *American Psychologist*, 45, 721-734.
- Auerbach, S. M. (1989). Stress management and coping research in the health care setting: An overview and methodological commentary. *Journal of Consulting and Clinical Psychology*, 57, 388-395.
- Band, E. B., & Weisz, J. R. (1990). Developmental differences in primary and secondary control coping and adjustment to juvenile diabetes. *Journal of Clinical Child Psychology*, 19, 150-158.
- Bauman, L. J., Drotar, D., Leventhal, J. M., Perrin, E. C., & Pless, I. B. (1997). A review of psychosocial interventions for children with chronic health conditions. *Pediatrics*, 100(2), 244-251.
- Bennett, D. S. (1994). Depression among children with chronic medical problems: A meta-analysis. *Journal of Pediatric Psychology*, 19, 149-169.
- Breslau, N. (1985). Depressive symptoms, major depression, and generalized anxiety: A comparison of self-reports on CES-D and results from diagnostic interviews. *Psychiatry Research*, 15, 219-229.
- Cadman, D., Boyle, M., Szatmari, P., & Offord, D. R. (1987). Chronic illness, disability, and mental and social well-being: Findings of the Ontario Child Health Study. *Pediatrics*, 79, 805-813.
- Chamberlain, P., & Reid, J. B. (1987). Parent observation and report of child symptoms. *Behavioral Assessment*, 9, 97-109.
- Compas, B. E., Malcarne, V. L., & Fondacaro, K. M. (1988). Coping with stressful events in older children and young adolescents. *Journal of Consulting and Clinical Psychology*, 56, 405-411.
- Coyne, J. C. (1997, August). In Alex J. Zautra (Chair), *Improving coping research: Raze the slum before any more building*. Symposium conducted at the annual meeting of the American Psychological Association, Chicago.
- Coyne, J. C., & Gottlieb, B. H. (1996). The mismeasure of coping by checklist. *Journal of Personality*, 64, 959-991.
- Coyne, J. C., Sonnega, J., & Carter, K. (1993, August). *Couples coping with congestive heart failure: A contextual approach*. Paper presented at the annual convention of the American Psychological Association, Toronto.
- Csikszentmihalyi, M., & Larson, R. (1984). *Being adolescent: Conflict and growth in the teenage years*. New York: Basic Books.
- Daniels, D., Moos, R. H., Billings, A. G., & Miller, J. J. (1987). Psychosocial risk and resistance factors among children with chronic illness, healthy siblings, and healthy controls. *Journal of Abnormal Child Psychology*, 15(2), 295-308.
- DiGirolamo, A. M., Quittner, A. L., Ackerman, V., & Stevens, J. (1997). Identification and assessment of ongoing stressors in adolescents with a chronic illness: An application of the behavior analytic model. *Journal of Clinical Child Psychology*, 26, 53-66.
- Drotar, D. (1994). Psychological research with pediatric conditions: If we specialize, can we generalize? *Journal of Pediatric Psychology*, 19, 403-414.
- Drotar, D. (1997). Relating parent and family functioning to the psychological adjustment of children with chronic health conditions: What have we learned? What do we need to know? *Journal of Pediatric Psychology*, 22, 149-165.
- D'Zurilla, T. J., & Goldfried, M. R. (1971). Problem solving and behavior modification. *Journal of Pediatric Psychology*, 78, 107-126.
- Espelage, D. L., & Quittner, A. L. (1999). Assessing social competence in women with and without eating disorders: A behavior analytic approach. Manuscript submitted for publication.
- Folkman, S., & Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior*, 50, 992-1003.

- Freund, A., Johnson, S. B., Silverstein, J., & Thomas, J. (1991). Assessing daily management of childhood diabetes using 24-hour recall interviews: Reliability and stability. *Health Psychology, 10*, 200–208.
- Gignac, M. A. M., & Gottlieb, B. H. (1996). Caregivers' appraisals of efficacy in coping with dementia. *Psychology and Aging, 11*, 214–225.
- Glueckauf, R. L. (1990). Program evaluation guidelines for the rehabilitation professional. In M. G. Eisenberg & R. Grzesiak (Eds.), *Advances in clinical rehabilitation* (Vol. 3, pp. 250–266). New York: Springer.
- Glueckauf, R. L. (1993). Use and misuse of assessment in rehabilitation: Getting back to the basics. In R. L. Glueckauf, L. B. Sechrest, G. R. Bond, & E. C. McDonel (Eds.), *Improving assessment in rehabilitation and health* (pp. 135–155). Newbury Park, CA: Sage.
- Glueckauf, R. L. (1996). Videosystem counseling for rural at-risk adolescents with epilepsy and their parents: An accessibility and outcome analysis (funded grant H133G60087). National Institute on Disability and Rehabilitation Research, USDE.
- Glueckauf, R. L. (in press). The Family and Disability Assessment System. In J. Touliatos & B.F. Perlmuter (Eds.), *Handbook of family measurement techniques* (Vol. 2). Newbury Park, CA: Sage.
- Glueckauf, R. L., Webb, P. M., Papandria-Long, M., Rasmussen, J. L., Markand, O., & Farlow, M. (1992). The family and disability assessment system: Consistency and accuracy of judgments across coders and measures. *Rehabilitation Psychology, 37*, 291–304.
- Goldberg, S., Morris, P., Simmons, R. J., Fowler, R. S., & Levison, H. (1990). Chronic illness in infancy and parenting stress: A comparison of three groups of parents. *Journal of Pediatric Psychology, 15*, 347–358.
- Goldfried, M. R., & D'Zurilla, T. J. (1969). A behavior-analytic model for assessing competence. In *Current Topics in Clinical Psychology, 1*, 151–195.
- Halstead, M., Johnson, S. B., & Cunningham, W. (1993). Measuring coping in adolescents: An application of the Ways of Coping Checklist. *Journal of Clinical Child Psychology, 22*, 337–344.
- Hauser, S. T. (1990). The study of families and chronic illness: Ways of coping and interacting. In G. Brody & I. Sigel (Eds.), *Methods of family research* (pp. 59–86). New York: Plenum Press.
- Hauser, S. T., DiPlacido, J., Jacobson, A. M., Paul, E., Bless, R., Milley, J., Lavori, P., Vieyra, M. S., Wolfsdorf, J. I., Herskowitz, R. D., Willett, J. B., Cole, C., & Wertlieb, D. (1993). The family and the onset of its youngster's insulin-dependent diabetes: Ways of coping. In R. E. Cole & D. Reiss (Eds.), *How do families cope with chronic illness?* (pp. 25–55). Hillsdale, NJ: Lawrence Erlbaum.
- Hobbs, N., Perrin, J. M., & Ireys, H. T. (1985). *Chronically ill children and their families*. San Francisco: Jossey-Bass.
- Hobfall, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44*, 513–524.
- Holmes, T., & Rahe, R. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research, 11*, 213–218.
- Johnson, S. B. (1995). Insulin-dependent diabetes mellitus in childhood. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (pp. 263–285). New York: Guilford.
- Johnstone, B., Frank F. G., Belar, C., Berk, S., Bieliauskas, L. A., Bigler, E. D., Caplan, B., Elliott, T. R., Glueckauf, R. L., Kaplan, R. M., Kreutzer, J. S., Mateer, C. A., Patterson, D., Puente, A. E., Richards, J. S., Rosenthal, M., Sherer, M., Shewchuk, R., Siegel, L. J., & Sweet, J. J. (1995). Psychology in health care: Future directions. *Professional Psychology: Research and Practice, 25*, 341–365.
- Kager, V. A., & Holden, E. W. (1992). Preliminary investigation of the direct and moderating effects of family and individual variables on the adjustment of children and adolescents with diabetes. *Journal of Pediatric Psychology, 17*(4), 491–502.
- Kanner, A. D., Coyne, J. C., Schaefer, C., & Lazarus, R. S. (1981). Comparisons of two models of stress management: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine, 4*, 1–39.
- Kazak, A. E., Segal-Andrews, A. M., & Johnson, K. (1995). Pediatric psychology research and practice: A family systems approach. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (pp. 84–104). New York: Guilford.
- Kovacs, M. (1985). The Children's Depression Inventory (CDI). *Psychopharmacology Bulletin, 21*, 995–998.
- LaGreca, A. M. (1994). Editorial: Assessment in pediatric psychology: What's a researcher to do? *Journal of Pediatric Psychology, 19*, 283–290.
- LaGreca, A. M., & Lemanek, K. L. (1996). Assessment as a process in pediatric psychology. *Journal of Pediatric Psychology, 21*, 137–151.

- Lavigne, J. V., & Faier-Routman, J. (1992). Psychological adjustment to pediatric physical disorders: A meta-analytic review. *Journal of Pediatric Psychology*, 17, 133–157.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Leviton, L. C. (1996). Integrating psychology and public health: Challenges and opportunities. *American Psychologist*, 51, 42–51.
- Lewis, J. M., Beavers, W. R., Gossett, J. T., & Phillips, V. A. (1976). *No single thread: Psychological health in family systems*. New York: Brunner/Mazel.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181–1209.
- Menaghan, E. G. (1983). Individual coping efforts and family studies: Conceptual and methodological issues. *Marriage and Family Review*, 6, 483–506.
- Meyerowitz, B. E., Heinrich, R. L., & Schag, C. C. (1983). A competency-based approach to coping with cancer. In T. G. Burish & L. A. Bradley (Eds.), *Coping with chronic disease: Research and applications* (pp. 137–158). New York: Academic Press.
- Montemayor, R. (1982). The relationship between parent-adolescent conflict and the amount of time adolescents spend alone and with parents and peers. *Child Development*, 53, 1512–1519.
- Moos, R. H., & Moos, B. S. (1981). *Family Environment Scale manual*. Palo Alto, CA: Consulting Psychologists Press.
- Mullins, L. L., Olson, R. A., Reyes, S., Bernardy, N., Huszti, H. C., & Volk, R. T. (1991). Risk and resistance factors in the adaptation of mothers of children with cystic fibrosis. *Journal of Pediatric Psychology*, 16(6), 701–715.
- Murch, R. L., & Cohen, L. H. (1989). Relationships among life stress, perceived family environment, and the psychological distress of spina bifida adolescents. *Journal of Pediatric Psychology*, 14(2), 193–214.
- O'Keefe, J., Quittner, A. L., & Melamed, B. (1996). Quality assurance and health outcomes for psychologists: A consensus report. In R. L. Glueckauf, R. G. Frank, G. R. Bond, & J. McGrew (Eds.), *Psychological practice in health care settings: Issues and new directions* (pp. 134–149). New York: Springer.
- Patterson, J. M. (1990). Family and health research in the 1980s: A family scientist's perspective. *Family Systems Medicine*, 8, 421–434.
- Pearlin, L. I., & Turner, H. A. (1987). The family as a context of the stress process. In S. V. Kasl & C. L. Cooper (Eds.), *Stress and health: Issues in research methodology* (pp. 143–166). New York: Wiley.
- Perrin, E. C., Stein, R. E., & Drotar, D. (1991). Cautions in using the Child Behavior Checklist: Observations based on research about children with a chronic illness. *Journal of Pediatric Psychology*, 16, 411–422.
- Pless, I. B., & Nolan, T. (1991). Revision, replication, and neglect. Research on maladjustment in chronic illness. *Journal of Child Psychology and Psychiatry*, 33, 347–365.
- Ptacek, J. T., Smith, R. E., Espe, K., & Raffety, B. (1994). Limited correspondence between daily coping reports and retrospective coping recall. *Psychological Assessment*, 6, 41–49.
- Quittner, A. L. (1992). Re-examining research on stress and social support: The importance of contextual factors. In A. M. La Greca, J. Wallander, L. Siegel, & C. Walker (Eds.), *Advances in pediatric psychology: Stress and coping with pediatric conditions* (pp. 85–115). New York: Guilford.
- Quittner, A. L. (1996). *Role-play inventory of situations and coping strategies (RISCS): Adolescent and parent-adolescent versions*. Bloomington: Indiana University.
- Quittner, A. L., & DiGirolamo, A. M. (1998). Family adaption to childhood disability and illness. In R. T. Amerman & J. V. Campo (Eds.), *Handbook of pediatric psychology and psychiatry* (pp. 70–102). Boston, MA: Allyn & Bacon Publishers.
- Quittner A. L., & Drotar, D. (1997). *Controlled trial of family interventions for cystic fibrosis* (funded grant proposal, National Institutes of Health). Bloomington: Indiana University.
- Quittner, A. L., & Espelage, D. (1997). Reliability and validity of a daily phone diary measure to assess daily activities and family interactions. Manuscript in press.
- Quittner, A. L., Espelage, D. L., Opiapari, L. C., Carter, B., Eid, N., & Eigen, H. (1998). Role strain in couples with and without a child with a chronic illness: Associations with marital satisfaction, intimacy, and daily mood. *Health Psychology*, 112–124.
- Quittner, A. L., Glueckauf, R. L., & Jackson, D. N. (1990). Chronic parenting stress: Moderating vs. mediating effects of social support. *Journal of Personality and Social Psychology*, 59, 1266–1278.
- Quittner, A. L., & Opiapari, L. C. (1994). Differential treatment of siblings: Interview and diary analyses comparing two family contexts. *Child Development*, 65, 800–814.

- Quittner, A. L., Opiplari, L. C., Regoli, M. J., Jacobsen, J., & Eigen, H. (1992). The impact of caregiving and role strain on family life: Comparisons between mothers of children with cystic fibrosis and matched controls. *Rehabilitation Psychology*, 37, 289–304.
- Quittner, A. L., Tolbert, V. E., Regoli, J. J., Orenstein, D., Hollingsworth, J. L., & Eigen, H. (1996). Development of the Role-Play Inventory of Situations and Coping Strategies (RISCS) for parents of children with cystic fibrosis. *Journal of Pediatric Psychology*, 21, 209–235.
- Quittner, A. L., Drotar, D., Ivvers-Landis, C., Seidner, D., Slocum, N., Jacobsen, J. (in press). Adherence to medical treatments in adolescents with cystic fibrosis: The development and evaluation of family-based interventions. In D. Drotar (Ed.), *Promoting adherence to medical treatment in childhood chronic illness: Concepts, methods, and interventions*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Raphael, K. G., Cloitre, M., & Dohrenwend, B. P. (1991). Problems with recall and misclassification with checklist measures of measuring stressful life events. *Health Psychology*, 10, 62–74.
- Reed, W. L., Richards, M. H., Moneta, G., Holmbeck, G., & Duckett, E. (1996). Changes in adolescents' daily interactions with their families from ages 10 to 18: Disengagement and transformation. *Developmental Psychology*, 32, 744–754.
- Reiss, D. (1981). *The family's construction of reality*. Cambridge, MA: Harvard University Press.
- Reiss, D., & Oliveri, M. E. (1980). Family paradigm and family coping: A proposal for linking the family's intrinsic adaptive capacities to its responses to stress. *Family Relations*, 29, 431–444.
- Reynolds, A. L., Johnson, S. B., & Silverstein, J. (1990). Assessing daily diabetes management by 24-hour recall interview: The validity of children's reports. *Journal of Pediatric Psychology*, 15, 493–509.
- Revenson, T. A., & Felton, B. J. (1989). Disability and coping as predictors of psychological adjustment to rheumatoid arthritis. *Journal of Consulting and Clinical Psychology*, 57, 344–348.
- Roberts, M. C., McNeal, R. E., Randall, C. J., & Roberts, J. C. (1996). A necessary reemphasis on integrating explicative research with the pragmatics of pediatric psychology. *Journal of Pediatric Psychology*, 21, 107–114.
- Robin, A. L. (1998). *ADHD in adolescents: Diagnosis and treatment*. New York: Guilford Press.
- Robson, A. L. (1997). Low birth weight and parenting stress during early childhood. *Journal of Pediatric Psychology*, 22(3), 297–311.
- Rock, D. L., Fordyce, W. E., Brockway, J. A., Bergman, J. A., & Spengler, D. M. (1984). Measuring functional impairment associated with pain: Psychometric analysis of an exploratory scoring protocol for activity pattern indicators. *Archives of Physical Medicine and Rehabilitation*, 65, 295–300.
- Rutter, M. (1990). Psychosocial resilience and protective mechanisms. In J. Rolf, A. S. Masten, D. Cicchetti, K. H. Neuchterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 181–214). New York: Cambridge University Press.
- Sandler, I. N., Wolchik, S., & Braver, S. (1988). The stressors of children's postdivorce environments. In S. A. Wolchik & P. Karoly (Eds.), *Children of divorce: Empirical perspectives on divorce* (pp. 185–232). New York: Gardner.
- Sechrist, L. B. (1993). Measurement in rehabilitation: From the beginning to what end? In R. L. Glueckauf, L. B. Sechrist, G. R. Bond, & E. C. McDonel (Eds.), *Improving assessment in rehabilitation and health* (pp. 135–155). Newbury Park, CA: Sage.
- Sechrist, L., McKnight, P., & McKnight, K. (1996). Calibration of measures for psychotherapy outcome studies. *American Psychologist*, 51, 1065–1071.
- Shiffman, S., Paty, J. A., Gnys, M., Kassel, J. A., & Hickcox, M. (1996). First lapses to smoking: Within-subjects analysis of real-time reports. *Journal of Consulting and Clinical Psychology*, 64, 366–379.
- Shonkoff, J. P., Hauser-Cram, P., Krauss, M. W., & Upshur, C. C. (1992). Development of infants with disabilities and their families: Implications for theory and service delivery. *Monographs of the Society for Research in Child Development*, Serial No. 230.
- Speechley, K. N., & Noh, S. (1992). Surviving childhood cancer, social support and parents' psychological adjustment. *Journal of Pediatric Psychology*, 17(1), 15–31.
- Spirito, A. (1996). Commentary: Pitfalls in the use of brief screening measures of coping. *Journal of Pediatric Psychiatry*, 21, 573–575.
- Spirito, A., Stark, L. J., & Tyc, V. L. (1994). Stressors and coping strategies described during hospitalization by chronically ill children. *Journal of Clinical Child Psychology*, 23, 314–322.
- Stark, L., & Quittner, A. L. (1996). *Behavioral treatment of weight gain in cystic fibrosis* (funded grant proposal, National Institutes of Health). Providence, RI: Brown University.
- Stone, A. A., Greenberg, M. A., Kennedy-Moore, E., & Newman, M. G. (1991). Self-report, situation-specific coping questionnaires: What are they measuring? *Journal of Personality and Social Psychology*, 61, 648–658.

- Stone, A. A., Kessler, R. C., & Haythornthwaite, J. A. (1991). Measuring daily events and experiences: Decisions for the researcher. *Journal of Personality*, 59, 575–607.
- Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*, 16, 199–202.
- Tinsley, H. E. A. (1992). Psychometric theory and counseling psychology research. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (pp. 37–70). New York: John Wiley & Sons.
- Turk, D. C., Sobel, H. J., & Follick, M. J., & Youkilis, H. (1980). A sequential criterion analysis for assessing coping with chronic illness. *Journal of Human Stress*, 6, 35–40.
- Varni, J. W., Wilcox, K. T., & Hanson, V. (1988). Mediating effects of family social support on child psychological adjustment in juvenile rheumatoid arthritis. *Health Psychology*, 7, 421–431.
- Vaux, A. (1985). Variations in social support associated with gender, ethnicity, and age. *Journal of Social Issues*, 41, 89–110.
- Wager, L., & Quittner, A. L. (April 1997). Development of an empirically based role-play measure to assess the problems and coping strategies of parents of adolescents with CF. Paper presented at the Sixth Florida Conference on Child Health Psychology, Gainesville, FL.
- Weiss, J. L., Marvin, R. S., & Pianta, R. C. (1997). Ethnographic detection and description of family strategies for child care: Applications to the study of cerebral palsy. *Journal of Pediatric Psychology*, 22, 263–278.
- 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, 688–701.
- Wertlieb, D., Weigel, C., & Feldstein, M. (1987). Measuring children's coping. *American Journal of Orthopsychiatry*, 57, 548–560.
- Wheaton, B. (1997). The nature of chronic stress. In B. H. Gottlieb (Ed.), *Coping with chronic stress* (pp. 43–73). New York: Plenum Press.
- Wysocki, T. (1993). Associations among teen-parent relationships, metabolic control, and adjustment to diabetes in adolescents. *Journal of Pediatric Psychology*, 18, 441–452.
- Wysocki, T., Harris, M. A., Greco, P., Bubb, J., Danda, C. L., Harvey, L. M., McDonnell, K., Taylor, A., & White, N. H. (in press). Randomized, controlled trial of behavior therapy for families of adolescents with insulin-dependent diabetes mellitus. *Journal of Pediatric Psychology*.
- Wysocki, T., White, N. H., Bubb, J., & Harris, M. A., & Greco, P. (1995). Family adaptation to diabetes: A model for intervention research. In J. L. Wallander & L. J. Siegel (Eds.), *Adolescent health problems: Behavioral perspectives* (pp. 289–304). New York: Guilford Press.

7

Qualitative Methods in Clinical Psychology

GLORIA L. KRAHN and DEBRA EISERT

Social science researchers are demonstrating a renewed interest in the use of qualitative methods to develop new knowledge and theory. In this chapter we explore this renewed interest by providing a brief overview of the historical preferences and precedents in research methods and by defining qualitative methods. We describe several of the more popular types of qualitative approaches, discuss criteria for evaluating adequacy of qualitative studies, demonstrate the application of qualitative methods in clinical child psychology, and finally, outline some current resources for training opportunities in qualitative methods.

HISTORICAL PREFERENCES AND PRECEDENTS IN THE USE OF QUALITATIVE METHODS

Psychology researchers are currently reevaluating the possible contributions of a variety of research methodologies, including an increase in the use of qualitative methods (Smith, Harre, & Van Langenhove, 1995a,b). Since the early 20th century, psychological researchers have relied primarily on quantitative methods based in the paradigm of logical positivism for their study of psychological phenomena. Positivism was founded on the philosophical position of Comte (1998) who, during the mid-19th century, argued that knowledge must be founded on observable information or “sense” data. He challenged psychology to abandon introspectionism and to base its findings on direct observation, sensory data, or statements that were deductively linked to those observations. Positivism continued to be a defining force on psychology research methods throughout the 20th

GLORIA L. KRAHN and DEBRA EISERT • Oregon Health Sciences University, Child Development and Rehabilitation Center, Portland, Oregon 97207-0574.

century. With a current decline in its influence, psychologists are considering expansion to other research approaches.

Criticisms of traditional quantitative methods have been voiced by diverse groups. For the past several decades, researchers in social psychology have lamented the narrowness of quantitative methods, with its reliance on laboratory studies, experimental design, and statistical analyses (Gergen, 1973, 1978; Harre & Secord, 1972; Shotter, 1975). The debate between advocates of quantitative and qualitative methods became contentious across numerous fields of inquiry, at times shedding more heat than light on the issues at hand. Some methodologists have advocated for more reasoned dialogue. Leading program evaluation researchers have maintained that quantitative and qualitative methods are not polar opposites (e.g., Reichardt & Cook, 1979) and that research methods should be selected on the basis of the research question rather than having the method determine the questions (e.g., Sechrest & Sidani, 1995).

Qualitative methods rely on data that are not easily quantified and focus on developing an understanding of the participants' perspectives of their experience and the context of the experience. They have been the traditional research methods in disciplines such as anthropology and sociology. Because of the richness of meaning that can be generated by qualitative methods, they have been growing in popularity in a number of disciplines related to psychology such as education (e.g., Reichardt & Cook, 1979; Patton, 1991), nursing (e.g., Habermann-Little, 1991; Inman, 1991), medicine (e.g., Blake, 1989; Waitzkin, 1990), rehabilitation (e.g., Huttlinger et al., 1992), and marketing (e.g., Stratton, 1991). They are also used more extensively by European psychologists (e.g., Banister, Burman, Parker, Taylor, & Tindall, 1994; Smith et al., 1995a) than North American psychologists. Researchers experienced in focus group methods or extensive unstructured observation are unlikely to be satisfied with single-method quantitative approaches to study complex social phenomena.

Researchers in child psychology already have a laudable history in qualitative methods. Early psychoanalysts developed their theories from case studies with clinical interviewing and disclosures in psychotherapy sessions. Piaget was known for his integration of multiple methods including clinical interviews, careful observations, and intensive case studies, which formed the basis for his hypothesis testing and theory development. His methods were not unlike the grounded theory approach of today. Bowlby (1969a,b) used primarily ethological observation methods in his well-known studies of attachment, separation, and loss. Systems theorist and social ecology researcher, Bronfenbrenner (1979), has been an advocate for the study of children in natural settings, strongly arguing for the study of the contexts within which children live such as their families, schools, neighborhoods, and cultures. Bruner (1990) continues to stress the importance of studying the ways in which people construct meaning in their unique cultures and argues for the use of a more diverse set of methodologies rather than sole reliance on the questions and methods of "yesterday's physics" (pp. xiii).

In spite of these influences, both historic and current, the general field of pediatric and clinical child psychology has followed a trend toward primary reliance on quantitative methods. While pediatric and child clinical researchers are familiar with qualitative methods such as clinical interviewing, direct observations, and intensive case studies, we are more accustomed to transforming these

data into linear, quantitative information. Survey reviews of publications in popular journals in this field verify that observation. In a recent review of publications in the *Journal of Clinical Child Psychology* from 1990 to 1993, we found only one article that was self-identified as qualitative in nature and only a few others that used some aspects of qualitative methods (e.g., Krahn, Hohn, & Kime, 1995). Similarly, a manual review of the *Journal of Pediatric Psychology* from 1979 into 1997 by the present authors identified no published articles that used solely qualitative methods.

Recent editorials in this field have recognized the need for more qualitative approaches. La Greca and Lemanek (1996) called for improved quality and focus on the assessments used in research: "In sum, measurement strategies that incorporate quantitative and qualitative approaches may help to capture the richness of the content areas that pediatric psychologists study" (p. 146). Structured and open-ended interviews were cited as especially appropriate for assessing the richness of the context and complex processes and to document meaning. Kazak (1997) opened the special issue of *Journal of Pediatric Psychology* on families and other systems by suggesting a need for new, more flexible approaches to the study of families. Suggestions included the need to frame questions from a family systems perspective, to aim to capture diversity in families, and to broaden the boundaries using social ecological or other contextual approaches. Qualitative methods may be particularly well suited to address these needs.

DEFINITION AND DESCRIPTIONS OF QUALITATIVE APPROACHES

Qualitative approaches are research designs, methods, and tools used to develop comprehension of a phenomenon through understanding its context and its meaning to the persons experiencing the phenomenon. Qualitative methods are often generically considered as those that produce data that are not readily reduced to numbers. Rather than reflecting a single homogeneous approach, qualitative approaches are diverse. Commonalities across qualitative approaches are (1) their roots in the phenomenological paradigm; (2) their concern with *verstehen*, or understanding human behavior from the individual's own frame of reference; (3) their holistic and inductive approach to develop knowledge; and (4) their basis in the data and resultant face validity and credibility (see Patton, 1990; Reichardt & Cook, 1979).

Questions, not hypotheses, serve as the beginning point for data collection in qualitative methods of research and program evaluation. Examples of questions for qualitative studies are: "How do parents experience being informed that their newborn child is disabled?" "What are the experiences of parents of children with conduct disorder?" and "What has occurred in disrupted adoptions, from the parents' perspective? from the child's?"

These questions are broad and open-ended, while at the same time addressing specific phenomena. In contrast to quantitative methods, qualitative methods are not intended and do not have the precision to confirm hypotheses or support theories, though they will often contribute to the development of theories and subsequent hypotheses (Needleman & Needleman, 1995). Instead, they help to develop understanding about the experiences that people encounter in their natu-

rally occurring contexts. They are largely inductive and develop understanding without imposing preexisting expectations or categories onto those experiences.

Qualitative data are typically represented as words or pictures whose meaning is not easily reduced to numbers. Researchers use a number of methods to collect qualitative data: (1) in-depth, open-ended interviews; (2) open-ended focus group discussions; (3) examination of documents or pictures; and (4) direct observation, including participant observation. As with all methods, these data collection methods have limitations that the researcher should be aware of in developing the research design. In interviews and focus groups, the interviewer and interviewee may hold different definitions of the issues being discussed. This discrepancy is aptly illustrated in the interchange, "Are you retarded? No, I'm Catholic," the title of Biklen and Moseley's (1988) qualitative study of persons with severe disabilities. These differences in definition of the construct are always a concern for self-report data, whether qualitative or quantitative in nature. Additionally, participants in unstructured or semistructured interviews and focus groups may report only on issues that are currently problematic and not report on issues that are related but resolved. This bias in reporting leads to a less complete description than the researcher may be seeking or may be interpreting. A safeguard against this is to continue to collect data until the data yield no new information. Such prolonged data collection will ultimately reveal a more comprehensive depiction of the phenomenon, but requires considerable resources. Patton (1990) recommends using observation in addition to participant report, and Greene and Caracelli (1997), among others, advocate for mixed-methods designs that include qualitative and quantitative methods.

Qualitative data collection methods require a higher level of involvement with respondents than do quantitative methods such as surveys. This involvement is particularly evident with participant observation when researchers join in family or community events to observe and record events and to reflect on their own reactions and experiences. Where objectivity and distance characterize data collection in survey research, in qualitative approaches a stance of empathetic neutrality by the researcher and the ability to separate observations from one's own reactions is important. Potential difficulties can occur in at least two areas. First, problems can arise with role confusion and possible role conflict in terms of the ethical responsibilities of researchers who become intimately familiar with the lives of their research participants. Anticipatory discussion of these conflicts and preestablished procedures for dealing with them can inform the research participants and can eliminate anxiety on the part of data collectors. Second, data collection is not "protected" by a structured protocol or established set of questions. This situation can lead to the data being vulnerable to the biases of the interviews through differential attention or probing in particular areas. Protections can be built into the data collection methods by thoroughly training interviewers or observers, including training in self-awareness of biases. This training can minimize interviewer biases from influencing the reporting or recording of data, while at the same time keeping attention focused on the key phenomena under study.

Sampling of participants in qualitative research is dependent on the purpose and scope of the study. In their study of parents' experiences in having a child with conduct disorder, Webster-Stratton and Spitzer (1996) followed a recommended qualitative research procedure of continuing to enroll new families until no new

information emerged; that is, continuing to sample participants until the conceptual framework was integrated, testable, and sufficiently explanatory of the phenomenon under study. Sampling in qualitative research is less concerned with randomness and more with purpose. One study may focus on homogeneous participants in order to describe a specific subgroup in depth such as newly homeless teens in a large urban setting, or it might focus on typical cases to develop a depiction of the usual types of participants, such as families receiving services in a pediatric oncology clinic and their customary concerns. In other studies, there may be a deliberate attempt to sample for maximum variability in order to describe the full range of possible experiences. Some studies deliberately include discrepant or extreme cases to gain the new information or insight that these participants can provide. While truncation of the data or elimination of outlier data points is seen as necessary with some quantitative analyses, qualitative approaches can accommodate heterogeneity among participants. The interested reader is referred to Patton (1990), who presents a more complete description of sampling approaches for different purposes.

Analyses of qualitative data is time intensive and is interactive with data collection. Often the data set is composed of hundreds of pages of narrative text. The researcher's task is one of sorting and sifting through this text, looking for commonalities of meaning and discrepancies across cases and contextual variables of the experience, in order to weave together the emerging theory about the phenomenon that then needs to be confirmed through further data collection. In approaches such as grounded theory (Glaser & Strauss, 1967; Strauss & Corbin, 1990), data collection is conducted conjointly with data analysis. Initially identified themes or constructs are checked against new data as they are collected and the constructs are revised accordingly, leading to a theory that has emerged from the data and is grounded in it.

Generalizability of the findings and theory resulting from qualitative methods is dependent on the purpose and procedures of the study. Because the intent of qualitative approaches is to understand a phenomenon from the frame of reference of the person living the experience, different experiences by different participants are valid and equally "true." Used in their traditional manner and depending on the sampling procedure, qualitative approaches allow for the possibility that alternative and different perspectives could be derived and that these alternative perspectives would be equally valid. These alternative perspectives may be particularly likely where sampling has included discrepant or extreme cases, and perhaps less likely in typical case sampling where the purpose of the study has been to identify the common essence of experience. This concern about generalizability can generate discomfort in researchers with a strongly quantitative perspective or whose conceptual framework is based on a belief in a singular "truth" (e.g., Clarke, 1992).

TYPES OF QUALITATIVE RESEARCH METHODS

A number of qualitative approaches are currently used and each one has a sufficiently distinct history and orientation to have retained unique identification. There are a number of valuable reference texts that provide an overview of these approaches that a researcher new to qualitative methods can consult (e.g., Lincoln

& Guba, 1985; Miles & Huberman, 1984, 1994; Patton, 1990; Strauss & Corbin, 1990). The following section briefly reviews four of the most popular approaches with applicability to child psychology.

Phenomenology

Many qualitative approaches share a common history in the broad field of phenomenology. Based in philosophy, phenomenology is the study of lived experiences through understanding the structure, essence, and context of the subjective experience of the individual (Beck, 1990). These experiences or phenomena can be variously defined as a setting such as a crowded shopping mall, an emotion such as "first love," or a situation such as inescapable conflict. Central to this approach is the study of how individuals perceive and make sense of an experience and determining the essence of the experience (Patton, 1990). Schroeder (1991) studied the psychological meaning of landscapes by examining arboretum users' memories and preferences for their favorite landscapes. Including photographs, thoughts, and feelings, he used quantitative and qualitative methods to identify dimensions of the psychological experiences of landscapes.

Grounded Theory

This qualitative method is growing in popularity within the social sciences. Grounded theory is based on the philosophical foundation of symbolic interaction and has become more accessible through the sociological work and writings of Strauss and his colleagues (Glaser & Strauss, 1967; Strauss, 1987; Strauss & Corbin, 1990). Its purpose is to develop new theory or refine existing theory directly from the raw data. Within the grounded theory approach, the theory is developed directly from the quotations of the participants or recorded notes of observers. Grounded theory method provides a prescribed process for analyzing qualitative data that may have been collected in a number of ways. With the growing use of focus groups for collecting participants' perspectives (e.g., Morgan, 1988), the grounded theory process can provide a systematic and credible method of analyzing those data. This process includes segmenting the narrative data into units, examining the units for their meaning, and grouping responses into as many categories as are needed to capture the content of the responses. New data are constantly compared with previous data and to the categories, allowing for revision of the categories concurrent with data collection. Relationships among these categories are examined for more abstract constructs. The theory emerges or is discovered by the researcher through the integration of these identified themes and constructs. Throughout this process the researcher pays careful attention to the trustworthiness of the data, the development of the response categories, and the interpretation of findings into theory.

Webster-Stratton and colleagues present a clearly articulated application of the grounded theory method to understanding the experiences of parents of children with conduct disorder (Spitzer, Webster-Stratton, & Hollinsworth, 1991; Webster-Stratton & Spitzer, 1996). In our own work, we applied a grounded theory method to understand parents' experiences of the process of being informed that their new-

born child was diagnosed with a significant disability (Krahn, Hallum, & Kime, 1993). As in all qualitative methods, the relationship between process and resulting theory is very important. Glaser and Strauss (1967) contend that the adequacy of the theory generated by the grounded theory approach cannot be separated from the process by which it is generated.

Ethnography

With its roots in anthropology and sociology, ethnography has been used to develop an understanding of ethnic and cultural groups. It is assumed that all groups of persons living together for an extended time develop their own culture. The term *culture* is used to refer to the integrated pattern of human behaviors that include thoughts, communications, actions, customs, beliefs, values, and institutions of a racial, ethnic, religious, or social group (Cross, Bazron, Dennis, & Isaacs, 1989). Ethnographers seek to become knowledgeable about cultural systems, components of the culture, and the rules that regulate the functioning of cultures. Ethnographic methods are based on extensive field work and participant observation in numerous settings and events. Such data may be supplemented by interviews with various members of the group under study. Ethnographic data are gathered as field notes, observations, tape recordings, written documents, or photographs. With increased attention to the cultural diversity within American society and efforts to increase cultural competence in provision of services (e.g., Anderson & Fenichel, 1989; Lynch & Hanson, 1992), ethnographic approaches can be a highly effective method to develop a better understanding of cultural patterns of children and families.

Burton (1991) provided a detailed and insightful ethnographic analysis of different cross-generational child care practices in an inner-city African-American community. Her findings elucidate the important roles that "uncle" figures and other members play in different family constellations and in diverse circumstances. Weiss, Marvin, and Pianta (1997) conducted an ethnographic study of families raising a child with cerebral palsy. Parents were interviewed about their experiences caring for the child and four distinct patterns of care emerged, three of which were considered adaptive in promoting continuity of the family. This study illustrates the importance of studying issues from the context of the family rather than making assumptions about what is adaptive. The paradigm of culture also has been applied increasingly to the study of organizations (e.g., Schaef & Fassel, 1988), and ethnographic approaches are helpful in understanding the unspoken rules and norms of behavior in different organizational settings.

Ecological Psychology

Based in an ecological systems perspective, ecological psychology views humans as interdependent with their settings and seeks to determine the relationships between the behavior of humans and their environments. Across different studies, these environments have been defined in terms that are physical (e.g., space, climate, surroundings) and interpersonal (e.g., who else is present? what do they say or do?). Ecological psychology allows for the inference of intentionality

in interpreting observed behavior. Data collection begins with detailed and comprehensive observations from a distanced, objective stance. These data are subsequently analyzed in terms of purposive, goal-directed action. As a method, ecological psychology gained prominence through the detailed, stream-of-behavior observations of Barker and Wright (1955) in describing one boy's day. This early work spawned a whole field of observational research in psychology that provided new insights into interpersonal interactions, particularly among family members (e.g., Patterson, 1975; Gottman, 1979). While initial and some subsequent studies have used extensive observations to describe behaviors and infer relationships among behaviors, many subsequent studies have become more quantitative, relying on preestablished observational coding systems and complex quantitative analyses to examine the interdependence among interactants' behaviors.

Bronfenbrenner (1979) applied an ecological psychology perspective to understand human behavior through consideration of its larger ecological context. Individuals are regarded as living and behaving within a nested series of ecological contexts. The ecological perspective of Bronfenbrenner has broadened the scope of contexts considered by researchers and policy analysts to understand human behavior, moving from examination of only immediate settings and interrelationships to consider the influences of broader community and societal factors. Fiese (1997) uses narrative data such as family stories to understand family adaptations and transitions within a transactional and ecological model. Family stories and rituals are described as contextual influences on the child's development, in regulating behavior and in establishing family norms.

CRITERIA FOR EVALUATING QUALITATIVE STUDIES

With the increased application of qualitative methods to psychology research has come the pragmatic need for criteria to evaluate the rigor of their research methods. A number of different sets of criteria have been proposed. We present two that are representative of the range of criteria proposed (e.g., Henwood & Pidgeon, 1992; Lincoln & Guba, 1985; Rennie, Phillips, Quartaro, 1988).

Lincoln and Guba

Lincoln and Guba (1985) developed a set of criteria that translate readily from quantitative approaches and provide a transition into qualitative methods for quantitatively trained researchers, such as most psychologists. They proposed a primary concept of *trustworthiness* that is broadly analogous to reliability and validity in quantitative studies. Trustworthiness is determined by the four criteria of credibility, transferability, dependability, and confirmability.

Credibility relates to confidence in the procedures and findings. The researcher wants to establish that the information being reported is the same as the informants' experiences. One way that credibility can be established is through triangulation, or the gathering of evidence from different points of view. This can be accomplished by interviewing several people about the same case, or by examining documents as alternative sources of information. A second way to establish cred-

ibility is to use member checking, which is the process of going back to the informants and determining whether the summarized information is consistent with the informants' experiences.

Transferability is concerned with the extent to which the findings are limited to their present context or are transferable to other groups or contexts. It is assessed by examining the characteristics of the study sample. For example, Bowman and Eisert (1997) used purposive sampling by choosing adolescent participants who shared the experience of a disrupted adoption. All participants were teenagers, so the information would not necessarily transfer to experiences of young children or adults.

Dependability is determined by the extent of agreement in categorizing or coding the data. It can be thought of as the reliability of the study's conclusions. The researchers need to demonstrate that the findings could be repeated using the same context and participants.

Confirmability is defined as the ability of the procedures to be formally audited and confirmed by independent review. Another researcher should be able to see the logic the researcher used to come to the final conclusions. Qualitative researchers should keep an audit trail using transcripts, field notes, observations, and a journal of personal reactions throughout the study so that the confirmability of the findings can be demonstrated.

Henwood and Pidgeon

An alternative set of criteria has been proposed by Henwood and Pidgeon (1992) that reflects the unique nature of qualitative methods more distinctly. Because qualitative approaches do not rest on a norm of objectivity that assumes independence of the known from the knower, these criteria move beyond measures that reduce observer bias. Henwood and Pidgeon propose the following seven criteria to assess the merit of qualitative methods:

- *Keeping close to the data* means being grounded in the data. Conceptual categories are summarized and labeled in ways that reflect their constituent data. An example would be to use comprehensive definitions that summarize why phenomena have been labeled in certain ways and that allow both researcher and reader to evaluate the fit of data with the label assigned.
- *Integrating theory at diverse levels of abstraction* is also drawn from grounded theory and places importance on how the data are synthesized. The goal is to ensure that the theory at all levels of abstraction is related to the data in meaningful and integrated ways. Readily apparent connections should be evident between the data and both lower and higher levels of abstractions.
- *Reflexivity* relates to a keen awareness of the research procedures used, the researcher's perspectives and biases, and the inevitable influence that the process of conducting the research has on the subject of study. Both feminist and cross-cultural researchers have advocated for increased self-consciousness by researchers. Procedures that address reflexivity include fully describing the researchers' attitudes and values or keeping a "reflexive journal" that

records the reflections of the researcher(s) own role and values, daily schedule of the study, and the methodological decisions made (see Lincoln & Guba, 1985).

- *Documentation* of the research process means generating records of the values and assumptions of the researchers, definitions of categories, initial concerns, sampling decisions, observations about the context, data collection procedures, and ideas about the quality of the data. Examples of documentation include research logs that provide a “paper trail” for external audit by other researchers and the records of classification decisions that are available with some computer programs for qualitative analyses of narrative data.
- *Theoretical sampling and negative case analysis* describes broad sampling that will develop or extend the theory and includes seeking cases that are apparent exceptions to the theory to aid in the conceptual development of the theory. An example would be to sample broadly among families and explore the meaning of the differences presented by families who are “outliers” or exceptions to the rule.
- *Sensitivity to negotiated realities* refers to developing a shared interpretation of the findings through dialogue with research participants; the resulting interpretation would be a negotiated joint reality between the perceptions of the participant and of the observations and interpretations of the researcher. In our work with families of newly diagnosed infants, the resultant negotiated realities challenged the researchers’ assumptions that the diagnostic information of significant disability was uniformly “bad news” (Krahn et al., 1993).
- *Transferability of findings* relates to how clearly the context of the study is recognized and has been described and to what other contexts the obtained findings might apply. In order to address the transferability of findings, the researcher must provide detailed information about the contextual features of the study.

There are obvious similarities across these two sets of criteria, while their differences reflect the differing assumptions of their underlying models of research. The criteria of Lincoln and Guba provide a ready translation from criteria for assessing rigor of quantitative approaches. As researchers become increasingly accustomed to qualitative methods and readers become more familiar with these approaches, we anticipate that the criteria based in qualitative approaches such as those of Henwood and Pidgeon will be used increasingly.

USE OF MULTIPLE METHODS

The earlier debate of qualitative methods versus quantitative methods is evolving to an appreciation of using multiple methods within the same study. Sechrest and Sidani (1995) argue that “good science is characterized by methodological pluralism” (p. 77) and describe the complementarity of quantitative and qualitative approaches to research questions. The practice of combining methods in this way has been labeled *triangulation*. In geographical surveying, triangulation refers to

the process of determining the position of an object by taking measurements of its position from two other points, the three points forming the corners of a triangle. In social science research, triangulation refers to the application of more than one method to the same research question to broaden the perspective and strengthen the findings of the study. This research approach has now been applied in numerous studies addressing diverse research questions, though not without some challenges in resolving discrepancies of different methods and data types as we have experienced ourselves (e.g., Krahn et al., 1995).

The focus of attention and debate has now shifted to the paradigms and premises of these new mixed-method designs. Combining multiple methods will intentionally or inadvertently also combine different and potentially conflicting paradigms that represent differing assumptions about the nature of the social phenomena under study. Greene and Caracelli (1997) define three levels of decision making to consider in mixed-method research design. These are the political level, or purpose of the study; the philosophical or paradigm level; and the technical or methods level. While mixed methods at the technical level can often strengthen a study, mixing methods at the paradigm level can pose greater difficulties.

Smith (1997) distinguishes paradigms from “crude mental models” and argues that it is these mental models that guide the actions of evaluators and researchers. She describes three types of mental models that researchers hold. The first model is based in the belief of a single reality that is separate from individuals’ interpretations of it. Model 1 presumes that definitive knowledge is possible. This has been the presumption of a positivist perspective and allows little, if any, role for qualitative approaches. The second model presumes a real world that is beyond the interpretations of any individual, but also one that cannot be studied free of individual perspective. In model 2, qualitative and quantitative components must be conducted separately and simultaneously to avoid the bias of interdependence. In the third mental model, the world is viewed as complex, contextually contingent and mediated by individual interpretation. In model 3, it is not possible to develop a definitive account of the phenomenon under study. The analysis of the data is the construction of the inquirer and social processes are as important as structural variables. This model is at the heart of qualitative “knowing” but appreciation or acceptance of this model is difficult for adherents of model 1. Delineation of these crude mental models is helpful in understanding the dialogue between qualitative and quantitative researchers.

APPLICATION OF QUALITATIVE METHODS IN CLINICAL CHILD PSYCHOLOGY

Examples of Grounded Theory

Several current examples of qualitative studies portray the application and usefulness of qualitative methods to clinical child psychology. Spitzer et al. (1991) studied the experiences and perceptions of parents who are raising a child with a conduct disorder. Through study of the process and content of a structured intervention program, they examined how parents change during the therapy process to cope with their children’s challenging behaviors. Using the constant comparison

method as an approach for developing a grounded theory, five phases were identified: acknowledging the problem, alternating despair and hope, "tempering the dream," "making the shoe fit," and effective coping. The authors suggest that, through studying the parents' experiences in therapy, researchers may be able to develop therapy process models, broaden outcome measures, and understand why families succeed or fail in therapy.

Bowman and Eisert (1995) studied the perceptions and experiences of children whose adoptions had been disrupted. Qualitative methods were applied that included multiple interviews and document analysis. Data collection used the topical oral biography, a specific form of life history in which children were asked to talk about a specific topic of their lives: the transitions involved in foster care placements, adoptions, and disruptions. Children were interviewed and their responses were coded thematically. Follow-up interviews were conducted after the initial coding to clarify ambiguous responses and to follow up on themes of interest. The children were interviewed until the researcher was satisfied that no new information was emerging. Document analysis and interviews with foster parents and protective service workers helped clarify the chronology of events in each child's life and described how children were supported during transitions.

The children's themes that emerged from the interviews included confusion and conflicted feelings, lack of choice and control, a sense of mourning and grief, and the need for permanence and belonging. Some children described a lack of belonging within the adoptive family or negative experiences such as being "the scapegoat"; others felt all was well in the family until they were suddenly asked to leave. Some children reported being misled by their adoptive families by promises that they would never have to leave or by promises to stay in contact after the disruption. The children voiced a need for more control in their lives and experienced these transitions from one family to another as unpredictable and unwarranted. They did not describe seeing their own behaviors or needs as being linked to decisions about them. All of the children expressed the feeling of not being wanted. One unanticipated finding was that for two of the participants who had experienced several adoptive disruptions, a subsequent placement worked well and the new parents stated that "there is nothing wrong with these children." Studies such as this can be useful when the researcher seeks to understand the life experience and perspectives of people encountering unique life circumstances. They can also help generate hypotheses, develop assessments, and capture the diversity of children's experiences.

Examples of Mixed Methods

There are numerous examples of mixed-method studies historically that combine open ended interviewing, thematic coding, and statistical analysis. In 1979, Frye and Simeonsson investigated the moral reasoning of typically developing young children and of adolescents and adults with mental retardation. They combined qualitative and quantitative scoring strategies to code children's responses to stories involving moral judgment. A Likert-type scale using happy and sad faces was used to quantify subjects' responses about how good or bad the child in the story was. Subjects were also asked to verbally justify their judgments and responses were coded qualitatively into categories such as ability, intent, and consequence. Similarly, Goldman, Whitney-Saltiel, Granger, and Rodin (1991) investi-

gated children's attributions of illness and health to determine if the attributes of adult concepts, such as causality, time line, consequence of illness, and cure dimensions, were present in children. Children's responses to open-ended questions such as "How do boys/girls get colds/fevers?" were solicited. Responses were scored using a Piagetian model based on children's construction of reality and concept of time. Data were statistically analyzed and results were interpreted as demonstrating the presence of immature but adultlike concepts in children's representations. These studies are typical of research that explores children's conceptual development.

Modifications of a grounded theory approach also are evident in past literature. Qualitative methods are used to obtain information that is subsequently coded into emerging themes. For example, Kistner et al. (1997) used open-ended clinical interviewing methods to determine children's intuitive theories about AIDS. Responses were then scored quantitatively to obtain information on accuracy of content and type of theory. La Greca et al. (1995) analyzed adolescents' quantitative and qualitative responses to questions about supportive behaviors in the care of diabetes. Some responses surprised the investigators. For example, some adolescents reported that "nagging" was a supportive behavior, if their friends reminded them to take their medication or test their insulin level. Nagging previously had been assumed to be a nonsupportive behavior. Krahn, Eisert, and Fifield (1990) developed a quantitative assessment of parent satisfaction with clinical services with two qualitative questions appended: "What did you like best?" and "What would you change?" Quantitative data were factor analyzed to generate four components of satisfaction and qualitative responses were categorized into these four factors. Of the narrative responses, 27% of the "What changed" and 5% of "liked best" could not be assigned to any of the factors and contributed new knowledge that was not revealed in the quantitative survey. These related to issues of inconvenience such as difficulty with parking, traveling long distances, and dissatisfaction with the physical facilities.

Role of Qualitative Methods in Developing New Measures

Qualitative approaches are often used as a first step in the development of new measures. Quittner et al. (1996) used structured interviews and daily diaries that were completed by families and health care professionals to develop detailed descriptions of problem situations in the care of children with cystic fibrosis. Responses were content analyzed and situations that were the most frequent or most difficult were incorporated into context-specific role-play vignettes that were audiotaped and used to elicit parents' coping strategies. This study illustrates how a behavior analytic model that emphasizes the situational nature of the stress and qualitative approaches that are used to obtain context-rich information can be combined successfully.

TRAINING PRACTICES AND RESOURCES IN QUALITATIVE METHODS

Despite the increasing use of qualitative methods in mixed method designs, clinical child psychologists still publish few studies that are primarily qualitative in nature. This situation likely relates to the biases of editors and researchers

toward quantitative studies and to the skills of psychologists who have little if any training in qualitative methods. There appear to be few opportunities for graduate students in clinical psychology to obtain training in qualitative methods. This lack of training is reflected in the nature of methods used in dissertation research. In a review of all clinical psychology doctoral abstracts for the years 1965 ($N = 249$) and 1985 ($N = 641$), Keeley, Shemberg, and Zaynor (1988) found only a modest increase in the use of nontraditional or nonquantitative methods from 2.5% in 1965 to 9.8% in 1985. These nontraditional methods included descriptive–interpretive studies (3.6%), survey methodologies (2.8%), library approaches (2.3%), and miscellaneous other approaches (1.1%).

The present authors conducted a telephone interview with a random sample of graduate clinical psychology training programs as presented on the Internet by the American Psychological Association. Of approximately 40 calls placed, only 9 programs were reached where a representative could speak to the training opportunities in the graduate program. Of these nine programs, only one reported offering a course within the psychology department in qualitative methods and one other program was considering such a course. Several other programs reported classes that might include limited coverage of qualitative approaches, for example, as part of an overview course. Two programs offered a course in program evaluation, another possible avenue for learning about qualitative approaches. Although this sample is small, it is also suggestive of the severe limitations in training on qualitative methods within clinical psychology training programs.

The field of psychology appears to lack research models for qualitative methods as well as mentors who are skilled in these methods. The discipline faces a significant challenge in its attempts to develop high-quality research using qualitative methods. At the same time, we witness that graduate psychology students are interested in learning more about qualitative methods and they are looking to other departments to gain such training. Clinical child and pediatric psychologists may wish to look to our past history of using these methods, as well as to current information and resources to increase and improve training opportunities.

Suggestions for Training

Researchers who seek an improved understanding of qualitative methods may become involved at different levels of complexity. At an introductory level, there are many sources of information about qualitative research that psychologists can pursue independently. Potential sources of information have been cited throughout this chapter. In addition, there are several texts in psychology that can be used as resources, including *Qualitative Methods in Psychology* (Banister et al., 1994), *Doing Qualitative Analysis in Psychology* (Hayes, 1997), *Rethinking Methods in Psychology* (Smith et al., 1995a), and *Qualitative Research in Psychology: Proceedings of the International Association for Qualitative Research in Social Science* (Ashworth, 1986). There are also emerging Internet sites on qualitative methods that can serve as resources for contacting other psychologists and learning about other materials as they are developed. For example, a website where researchers can try a demonstration of qualitative data analysis using the Nudist program is currently available at www.scolari.com. Other websites provide a discussion format for a qualitative methods group (www.psyc.nott.ac.uk/qm.html), description of an advanced qualitative methods course in psychology (psychserv.psych.ucalgary.ca).

ca/CourseNotes/PSYC503/w97.html#RTFToC1), and an Internet guide to narrative psychology (web.lemoyne.edu/~hevern/narpsych.html).

A second level of involvement is for the discipline to seek qualified researchers, perhaps from other disciplines, to offer training workshops. These workshops could provide an overview of methods and serve to connect interested researchers with qualified mentors. Workshops could be specific to the types of qualitative analyses that are pertinent to specific research questions and kinds of available data. These workshops could be offered through continuing education venues such as regional and national meetings of the American Psychological Association in order to attract a wide audience and to indicate a professional acceptance of these methods by the discipline.

Pediatric and clinical child psychologists who wish to educate themselves further should consider enrolling in official course work. Many major universities offer training in qualitative methods in other departments. For example, at the University of Oregon, courses in qualitative methods are offered in the departments of sociology, anthropology, and special education. At medical universities, the school of nursing may offer such course work.

At the most advanced level, where researchers are considering ongoing use of qualitative research, completion of training through course work or workshops and work with mentors or collaborators with expertise in qualitative methods is necessary. Qualitative research is not a unitary field and therefore not easily mastered. The pursuit of qualitative research by pediatric and clinical child psychologists will provide an additional set of methodological tools that are available for appropriate research questions and for data that cannot easily be quantified.

FUTURE DIRECTIONS FOR RESEARCH USING QUALITATIVE METHODS IN PEDIATRIC AND CLINICAL CHILD PSYCHOLOGY

Over a decade ago, Wicker (1985, 1989) invited psychologists to "get out of our conceptual ruts" and to engage in "substantive theorizing" in many of the ways we now see occurring through consideration of context and reported experiences in research. As the field of psychology reconsiders its methodological assumptions and moves to expand its accepted repertoire of research approaches, researchers in pediatric and clinical child psychology have the opportunity to use new approaches to study traditional and emerging questions. Qualitative approaches are applicable across a broad array of research questions and are likely to be used as one component of a multimethod approach. A popular example is the use of focus groups to identify relevant dimensions for subsequent questionnaire surveys. Similarly, some researchers have interpreted the findings of quantitative studies through the subsequent use of key informant interviews or focus groups. An ecological perspective serves as a framework for considering future directions and the unique contributions qualitative approaches can make to research at the individual, family, community, and larger systems levels.

Individual Level Applications

Qualitative methods are particularly valuable in the study of low-incidence situations or disorders where the small sample size makes quantitative study

difficult. The resulting study can provide in-depth understanding of the experience, the context, and many implications of the disorder that would not be included in the scope of a quantitative study. An example is Todis's (1996) study of students with significant disabilities using assistive technology in educational settings. Inconsistent use of assistive devices such as speech simulators is a common source of frustration for professionals. Todis's findings provide insight into when and why students and their families and teachers do not use devices regularly across settings. Such understanding can serve as the basis for modifying the devices and training for their use and for changing the expectations and approaches of therapists and teachers.

Qualitative approaches are uniquely capable of developing rich insight into intensely felt personal experiences such as life-threatening illnesses, living with a disability, or experiencing domestic violence or the death of a parent. While our professional literature contains limited external reports of such experiences, we have less information about the personal experience of such situations and subsequent influence on development. One example is Ronai's (1997) account of being raised by a mother with mental retardation. Within an ethnographic approach she uses a layered account of research data, personal narration of lived experience, and social theory to portray vividly and emotionally her process of understanding her own experiences and feelings.

Family and Community-Level Applications

Efforts at understanding the influence of family composition and interactions on child development have only been partially successful, despite a large quantitative literature. Inclusion of qualitative methods within a multimethod design can return the researcher to a more connected understanding of the experience under study. Qualitative approaches provide a means to consider the complex interplay of interpersonal interactions within a family or community group. As illustrated by Ronai's (1997) ethnographic account, the naturally occurring host of interpersonal variables can be incorporated into the study in a way that would be unworkable in a quantitative design with its need to control for each additional variable or measure it with a resulting need for a larger sample size. The qualitative approach allowed for the incorporation of other contextual variables (e.g., sexually deviant father; empathic social worker) to describe the richly textured experiences of her childhood and adult life.

Psychology is demonstrating increased attention to cultural diversity and the reflections of culture in behaviors and beliefs. Past research practices have unreflectively applied assessment instruments and procedures to persons of varying ethnic backgrounds, without an appreciation of the possibly different cultural interpretations or meanings of the phenomenon under study or data collection practices. Issues that are critical to pediatric and child psychologists can vary across cultural communities. These include:

1. The definition and meaning of family (who is regarded by the family as being part of the family? as responsible for child rearing? who should be included in data collection?).
2. Cultural values (to what extent does this culture value interdependence and community relative to independence and individualism?).

3. Emphasis on interpersonal control and compliance (what practices are used to ensure social compliance? who is seen as responsible for child compliance?).
4. Data collection methods (how will families respond to questionnaires or interviews? are focus groups with similar members more effective for information disclosure?).

Qualitative methods provide a means of exploring the cultural meanings of events and practices that can support subsequent quantitative study.

Societal and Larger Systems Applications

Psychology has intermittently recognized the influences of larger systems on the human experience. Yet, the influence of societal or larger systems context on human experience and behavior is often difficult to recognize and understand. Not until we pull our nose away from the glass do we recognize that we have been looking at an experience through only one window. Abrupt changes in time or differences across geopolitical areas provide natural opportunities to examine these larger systems influences.

The current and rapid change to managed health care in the United States is one such opportunity. As changes are brought about in the financing mechanisms for health care delivery, new market forces impact on the experiences of families and professionals and on the relationship between professional and family. Even alternative managed care payment formats yield differences in perceived access to care and the nature of care received. Focus groups have proved to be a valuable method to understand these influences on the experiences of particularly vulnerable populations such as children with special health care needs (e.g., Krahm & Hartzell, 1996). These qualitative data have helped define significant dimensions for subsequent quantitative surveys.

Finally, and coming full circle in this chapter, even the reluctance of North American psychologists relative to European psychologists to embrace qualitative approaches has been attributed to their differing sociopolitical contexts. According to Danziger (1979), American psychologists have been held accountable to a different audience, namely the business community and politicians who have more influence over research funds and academic appointments in this country than in others. Only recently in the United States have we seen a modest departure from adherence to the objective and physical science model to include research funding availability for qualitative methods and alternative health care practices. Psychology may soon demonstrate broader acceptance of qualitative approaches for research; it has a history and set of resources that psychologists can draw on to develop their own expertise and effectively use qualitative research approaches.

REFERENCES

- Anderson, P. P., & Fenichel, E. S. (1989). *Serving culturally diverse families of infants and toddlers with disabilities*. Arlington, VA: National Center for Clinical Infant Programs.
- Ashworth, P. D., Giorgi, A., & deKonig, A. J. J. (Eds.). (1986). Qualitative research in psychology: Proceedings of the International Association for Qualitative Research in Social Sciences. Pittsburgh, PA: Duquesne University Press.

- Banister, P., Burman, E., Parker, I., Taylor, M., & Tindall, C. (Eds.). (1994). *Qualitative methods in psychology: A research guide*. Buckingham, England: Open University Press.
- Barker, R. G., & Wright, H. F. (1955). *Midwest and its children*. New York: Harper & Row.
- Beck, C. T. (1990). Qualitative research: Methodologies and use in pediatric nursing. *Issues in Comprehensive Pediatric Nursing*, 13, 193-201.
- Biklen, S. K., & Moseley, C. R. (1988). "Are you retarded?" "No, I'm Catholic": Qualitative methods in the study of people with severe handicaps. *Journal of the Association for Persons with Severe Handicaps*, 13(3), 155-162.
- Blake, R. L., Jr. (1989). Integrating quantitative and qualitative methods in family research. *Family Systems Medicine*, 7, 411-427.
- Bowlby, J. (1969a). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.
- Bowlby, J. (1969b). *Attachment and loss: Vol. 2. Separation: Anxiety and anger*. New York: Basic Books.
- Bowman, B. B. (1995). Children's perspectives of disrupted adoption: A qualitative study of lives within the child welfare system. Unpublished doctoral thesis, University of Oregon.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Burton, L. (1991). Intergenerational patterns which influence the nurturing domain: Infants, parents and grandparents. Presentation at the National Center for Clinical Infant Programs conference, Washington, DC, December 7, 1991.
- Clarke, L. (1992). Qualitative research: Meaning and language. *Journal of Advanced Nursing*, 17, 243-252.
- Comte, A. (1998). *Early political writings* (edited and translated by H. S. Jones). Cambridge, NY: Cambridge University Press.
- Cross, T. L., Bazron, B. J., Dennis, K. W., & Isaacs, M. R. (1989). *Towards a culturally competent system of care, Vol. 1*. Available from the Georgetown University Child Development Center, 3307 M Street, NW, Washington, DC.
- Danzinger, K. (1979). The social origins of modern psychology. In A. Buss (Ed.), *Psychology in social context*. New York: Irvington.
- Fiese, B. H. (1997). Family context in pediatric psychology from a transactional perspective: Family rituals and stories as examples. *Journal of Pediatric Psychology*, 22(2), 183-196.
- Frye, H., & Simeonsson, R. J. (1979). Quantitative and qualitative analyses of moral reasoning in children, adolescents, and adults of similar mental age. *Journal of Pediatric Psychology*, 1, 197-210.
- Gergen, K. (1973). Social psychology as history. *Journal of Personality and Social Psychology*, 26, 309-320.
- Gergen, K. (1978). Experimentation in social psychology. *European Journal of Social Psychology*, 8, 507-527.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine.
- Goldman, S. L., Whitney-Saltiel, D., Granger, J., & Rodin, J. (1991). Children's representations of "everyday" aspects of health and illness. *Journal of Pediatric Psychology*, 16, 747-766.
- Gottman, J. M. (1979). *Marital interaction: Experimental investigations*. New York: Academic Press.
- Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation*, 74, 5-17.
- Habermann-Little, B. (1991). Qualitative research methodologies: An overview. *Journal of Neuroscience Nursing*, 23, 188-190.
- Harre, R., & Secord, P. F. (1972). *The explanation of social behavior*. Oxford: Blackwell.
- Hayes, N. (Ed.). (1997). *Doing qualitative research in psychology*. Hove, England: Psychology Press/Erlbaum.
- Henwood, K. L., & Pidgeon, N. F. (1992). Qualitative research and psychological theorizing. *British Journal of Psychology*, 83, 97-111.
- Huttlinger, K., Krefting, L., Drevdahl, D., Tree, P., Baca, E., & Benally, A. (1992). "Doing battle": A metaphorical analysis of diabetes mellitus among Navajo people. *American Journal of Occupational Therapy*, 46, 706-712.
- Inman, C. E. (1991). Analyzed interaction in a children's oncology clinic: The child's view and parent's opinion of the effect of medical encounters. *Journal of Advanced Nursing*, 16, 782-793.
- Kazak, A. (1997). A contextual family/systems approach to pediatric psychology: Introduction to the special issue. *Journal of Pediatric Psychology*, 22, 141-148.

- Keeley, S. M., Shemberg, K. M., & Zaynor, L. (1988). Dissertation research in clinical psychology: Beyond positivism? *Professional Psychology: Research and Practice*, 19, 216–222.
- Kistner, J., Eberstein, I. W., Balthazar, M., Castro, R., Foster, K., Osborne, M., Sly, D., & Quadagno, D. (1997). Assessing children's conceptions of AIDS. *Journal of Pediatric Psychology*, 21, 269–282.
- Krahn, G. L., Eisert, D., & Fifield, B. (1990). Obtaining parental perceptions of the quality of services for children with special health needs. *Journal of Pediatric Psychology*, 15, 761–774.
- Krahn, G. L., Hallum, A., & Kime, C. (1993). Are there good ways to give "bad news"? *Pediatrics*, 91, 578–582.
- Krahn, G. L., & Hartzell, M. S. (1996). Children with special health care needs and managed care in Oregon. Presentation at the American Public Health Association, New York, November 1996.
- Krahn, G. L., Hohn, M. F., & Kime, C. (1995). Incorporating qualitative approaches into clinical child psychology research. *Journal of Clinical Child Psychology*, 24, 204–213.
- LaGreca, A. M., Auslander, W. F., Greco, P., Spetter, D., Fisher, E. G., Jr., & Santiago, J. V. (1995). I get by with a little help from my family and friends: Adolescents' support for diabetes care. *Journal of Pediatric Psychology*, 20, 449–476.
- La Greca, A. M., & Lemanek, K. L. (1996). Editorial: Assessment as a process in pediatric psychology. *Journal of Pediatric Psychology*, 21, 137–151.
- Lincoln, Y. S., & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Lynch, E. W., & Hanson, M. J. (1992). *Developing cross-cultural competence: A guide for working with young children and their families*. Baltimore: Brookes.
- Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis*. Beverly Hills, CA: Sage.
- Miles, M. B., & Huberman, A.M. (1994). *Qualitative analysis: An expanded sourcebook* (2nd ed.). Beverly Hills, CA: Sage.
- Morgan, D. L. (1988). *Focus groups as qualitative research*. Beverly Hills, CA: Sage.
- Needleman, C., & Needleman, M. L. (1995). Qualitative methods for intervention research. *American Journal of Industrial Medicine*, 29, 329–337.
- Patterson, G. R. (1975). *Families: Applications of social learning to family life* (rev. ed.). Champaign, IL: Research Press.
- Patton, M. J. (1991). Qualitative research on college students: Philosophical and methodological comparisons with the quantitative approach. *Journal of College Student Development*, 32, 389–396.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newberry Park, CA: Sage.
- Quittner, A. L., Tolbert, V. E., Regoli, M. J., Orenstein, D. M., Hollingsworth, J. L., & Eigen, H. (1996). Development of the Role-play Inventory of Situations and Coping Strategies for parents of children with cystic fibrosis. *Journal of Pediatric Psychology*, 21, 209–235.
- Reichardt, C. S., & Cook, T. D. (1979). Beyond qualitative versus quantitative methods. In T. D. Cook & C. S. Reichardt (Eds.), *Sage research progress series in evaluation: Vol.1. Qualitative and quantitative methods in evaluation research* (pp. 7–32). Beverly Hills, CA: Sage.
- Rennie, D. L., Phillips, J. R., & Quartaro, G. K. (1988). Grounded theory: A promising approach to conceptualization in psychology. *Canadian Psychology*, 29, 139–150.
- Ronai, C. R. (1997). On loving and hating my mentally retarded mother. *Mental Retardation*, 35, 417–432.
- Schaef, A. W., & Fassel, D. (1988). *The addictive organization*. San Francisco, CA: Harper & Row.
- Schroeder, H. W. (1991). Preference and meaning of arboretum landscapes: Combining quantitative and qualitative data. *Journal of Environmental Psychology*, 11, 231–248.
- Sechrest, L., & Sidani, S. (1995). Quantitative and qualitative methods: Is there an alternative? *Evaluation and Program Planning*, 18, 77–87.
- Shotter, J. (1975). *Images of man in psychological research*. London: Methuen.
- Smith, J., Harre, R., & Van Langenhove (Eds.). (1995a). *Rethinking methods in psychology*. London: Sage.
- Smith, J., Harre, R., & Van Langenhove (Eds.). (1995b). *Rethinking psychology*. London: Sage.
- Smith, M. L. (1997). Mixing and matching: Methods and models. *New Directions for Evaluation*, 74, 73–85.
- Spitzer, A., Webster-Stratton, C., & Hollinsworth, T. (1991). Coping with conduct-problem children: Parents gaining knowledge and control. *Journal of Clinical Child Psychology*, 20, 413–427.
- Stratton, P. (1991). Attributions, baseball and consumer behaviour. *Journal of the Market Research Society*, 33, 163–178.
- Strauss, A. K. (1987). *Qualitative analysis for social scientists*. Cambridge: Cambridge University Press.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research*. Beverly Hills, CA: Sage.
- Todis, B. (1996). Tools for the task? Perspectives on assistive technology in educational settings. *Journal of Special Education Technology*, 13, 49–61.

- Waitzkin, H. (1990). On studying the discourse of medical encounters. *Medical Care*, 28, 473–487.
- Webster-Stratton, C., & Spitzer, A. (1996). Parenting a young child with conduct problems: New insights using qualitative methods. In T. H. Ollendick & F. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 18, pp. 1–62). New York: Plenum Press.
- Weiss, K. L., Marvin, R. S., & Pianta, R. C. (1997). Ethnographic detection and description of family strategies for child care: Applications to the study of cerebral palsy. *Journal of Pediatric Psychology*, 22, 263–278.
- Wicker, A. W. (1985). Getting out of our conceptual ruts: Strategies for expanding conceptual frameworks. *American Psychologist*, 40, 1094–1103.
- Wicker, A. W. (1989). Substantive theorizing. *American Journal of Community Psychology*, 17, 531–547.

8

Understanding Cultural and Ethnic Influences in Research with Child Clinical and Pediatric Psychology Populations

NATALIE WALDERS and DENNIS DROTAR

RATIONALE FOR CONDUCTING RESEARCH WITH ETHNICALLY DIVERSE POPULATIONS

It is becoming increasingly necessary for child clinical and pediatric psychologists to possess the skills and knowledge necessary to conduct research and clinical work with ethnically diverse populations (Hall, 1997). Despite increasing demographic diversity within the United States, psychological research has been largely restricted to mainstream, homogeneous samples (Foster & Martinez, 1995). Consequently, the knowledge base concerning both normative and problematic developmental processes within ethnic minority child clinical and pediatric populations remains limited. Graduate and professional level training opportunities for conducting ethnically diverse research are also limited in scope and availability (Ponterotto & Cases, 1987).

The purpose of the present chapter is to document the critical need for an increased focus on ethnic minority children and families in child clinical and pediatric psychological research, to review and examine the specific challenges

NATALIE WALDERS • Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123. **DENNIS DROTAR** • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

associated with research in this area, and to offer recommendations on how to overcome these problems. Finally, we will present models of successful clinical child and pediatric research with ethnic minority children and families and describe implications for training.

Demographic Shifts in the US Population

The demographic landscape of the US population has diversified at a dramatic pace in recent years, and future projections anticipate a continued increase in the number of ethnic minorities in this country. According to the US Bureau of the Census (1995), by the year 2000, it is projected that approximately one third of people living in the United States will be ethnic minorities. By the year 2050, the country is expected to grow increasingly diverse, such that only 50% of the population will be identified as white. As the composition of the United States continues to diversify, the need for culturally informed approaches to psychological research and practice will become more urgent (Bernal & Castro, 1994). According to Census projections, ethnic minority representation will expand most dramatically within child and adolescent populations. By 2030, non-Hispanic whites under age 18 will comprise less than half of the population, while, 75% of the US population age 65 will remain non-Hispanic white (US Bureau of the Census, 1995). Minority children and adolescents are at a disproportionately elevated risk for a number of health problems. They also are exposed to a constellation of mental and physical health risk factors at a much higher rate than white children, most notably, those associated with poverty (Kawaga-Singer, Katz, Taylor, & Vanderryn, 1996).

Limited Research Database on Ethnic Minority Issues within Pediatric and Clinical Child Psychology

Numerous literature reviews have demonstrated a neglect of ethnic minority-focused research topics and have indicated that ethnic minorities have been largely overlooked in a majority of research studies (Iwamasa & Smith, 1996; Ponterotto, 1988; Reid & Kelley, 1994). For example, Graham's (1992) review of six American Psychological Association journals for empirical studies on African-Americans over two decades found a significant decline in the number of studies focusing on African-American issues over the 20 year time span. Throughout the time period, only 3.6% of published studies within these journals focused on African-Americans. In a review of the school psychology literature, Rogers-Wiese (1992) found a limited number of studies focusing on issues facing ethnic minority children. The majority of the few published studies examined psychoeducational assessment for African-American children. A brief review of articles published in the *Journal of Pediatric Psychology* over the course of 1 year indicated that approximately 25% of participants in studies published during 1996 were from ethnic minority backgrounds. However, it should be noted that many published articles did not incorporate descriptive information on samples, making it difficult to obtain an exact overview of demographic data. Overall, one published study focused exclusively on an ethnic minority pediatric health concern, examining hope, coping, and adjustment in African-American children with sickle cell disease (Lewis & Kliewer, 1996).

In response to the limited research database on ethnic minorities within psychology, several independent journals have been initiated to provide an established forum for these topics, for example, *Journal of Black Studies* and *Journal of Multicultural Counseling and Development*. However, to our knowledge, no journals have been launched to address culture and/or ethnicity with specific reference to child clinical and pediatric psychology.

Consequences of Neglecting Ethnic Minority Issues in Child Clinical and Pediatric Research

The neglect of cultural diversity in research has led to several specific negative consequences for the knowledge base in child clinical and pediatric psychology. The focus on homogeneous populations in major research studies has contributed to inaccurate generalizations and misconceptions concerning ethnic minority children and families (Kawaga-Singer, et al., 1996; Ogbu, 1985). For example, McKenry, Everett, Ramseur, and Carter (1989) argued that a lack of diversity in research contributed to problematic and premature conclusions that differences between African-American and white adolescent experiences represented African-American developmental deficiencies and pathologies. In addition, Stanton and colleagues (1997) cautioned us that longitudinal research on the development of youth risk behaviors largely has been conducted with white, middle-class samples, thus limiting the applicability of intervention research for ethnically diverse populations. Cross-cultural developmental psychologists also have noted the problems associated with relying on homogenous samples to formulate normative developmental theory (Rogoff & Morelli, 1994).

Overall, ethnic minorities are underrepresented in mental health research studies, which limits the development, implementation, and evaluation of culturally inclusive treatment and intervention programs (Kato & Mann, 1996). Without the data necessary to shape psychotherapeutic interventions from a culturally informed perspective, individual clinicians who work with children and families are at a disadvantage. Moreover, broad-based clinical and preventative intervention programs may be unable to best impact diverse communities. In addition, the concentration on mainstream samples prevents minority children and their families from accessing the potential benefits associated with participation in research trials. Finally, the lack of a focus on diversity in intervention and outcome research limits psychologists' ability to successfully serve as culturally informed policy advocates on behalf of minority children and families.

DESIGN AND IMPLEMENTATION CHALLENGES IN ETHNIC MINORITY-FOCUSED CHILD CLINICAL-PEDIATRIC PSYCHOLOGY RESEARCH

Research Design Decisions: Issues of Terminology

One of the main difficulties in conducting inclusive and informed research involves the lack of consensus concerning appropriate definitions and applications of terms such as race, ethnicity, and culture (Foster & Martinez, 1995). Researchers have tended to use these terms interchangeably, paying limited attention to their

specific meanings and various interpretations (Alvidrez, Azocar, & Miranda, 1996; Betancourt & Lopez, 1993; Beutler, Brown, Crothers, Booker, & Seabrook, 1996). In addition, researchers often have neglected to address potentially confounding factors of socioeconomic status, sex differences, and geographic origin when attempting to define and classify diverse samples. Consequently, we begin with a clarification of terminology.

Race

Traditionally, race has been conceptualized as a biological marker distinguishing a group of individuals with similar genetic makeup and geographic origin and shared visible physical traits (Zuckerman, 1990). The use of race as a biologically linked variable has been challenged by many scientific disciplines, including psychology. Critics of the use of race as a valid categorization have argued convincingly that racial delineations are arbitrary in nature and lack scientific support (Beutler et al., 1996; LaVeist, 1994; Zuckerman, 1990). Racial categories have been questioned for their biological and genetic validity as well as for their societal interpretations. The use of race as an identifying variable has overshadowed within-group variability and has been criticized as an inaccurate and imprecise means of defining groups of individuals (Jenkins & Parron, 1995).

The research implications of utilizing race as a grouping variable also have been emphasized: "Its [race's] varied meanings and connotations create confusion, incite controversy and partisanship, and limit research design and theory building" (Yee, Fairchild, Weizmann, & Wyatt, 1993, p. 1134). According to critics of the application of race in psychological research, the term lacks empirical support and specificity necessary for use in such contexts. Employing race as a categorical variable for research has led to an underestimation of within-group variation and a confounding of biological, behavioral, and social factors (Wyatt, 1991; Zuckerman, 1990).

Culture

Culture has been defined in a number of nonspecific overarching ways. It is typically employed as an umbrella term to describe a wide range of variables (Triandis & Brislin, 1984). Carter (1995) defines culture from a psychological perspective as the process of communicating and transmitting systems of language, knowledge, attitudes, and behaviors between generations. Culture typically incorporates ethnic identity, physical characteristics, and external societal influences. From a developmental perspective, culture may be conceived of as the broad "shared elements that provide the standards for perceiving, believing, evaluating, communicating, and acting" (Triandis, 1996, p. 408).

Recommendations for Terminology: Ethnicity and Specificity

When such data are available, utilizing ethnicity as a descriptor variable is the preferred alternative for classification because it involves features of race and culture and serves as a more comprehensive means of identifying an individual's background. Ethnic groupings are determined by a range of criteria, including such

indicators as language background, geographic location, and acculturative status. As highlighted by Jenkins and Parron (1995), employing ethnicity rather than race as a categorization variable offers the opportunity to explore potentially important differences between subpopulation groups. For example, reanalyzing national data on Hispanic adolescents according to detailed ethnic variations revealed significant behavioral differences between Puerto Rican, Cuban-American, and Mexican-American youth (Sokol-Katz & Ulbrich, 1992; Jenkins & Parron, 1995). An extended discussion of the issues and options surrounding classification of ethnicity for research applications may be found in a recent document issued by the National Institutes of Health (1995).

In order to avoid problematic classification of participants in research, several useful specific recommendations have been made for assessing ethnicity within child clinical or pediatric research. First, due to the lack of a universally accepted and applied definition of the variables comprising ethnic status, it is highly desirable for researchers to explicitly state which features (e.g., family background, language of origin, place of birth) were used when defining ethnicity (Beutler et al., 1996). In addition, it is important to describe the ethnic breakdown of a sample regardless of whether ethnicity was used as a variable for purposes of data analysis.

Researchers also have recommended going beyond a simplistic "box checking" procedure when classifying study samples. For example, inclusion of factors such as language, acculturation, and community composition has been proposed for sample description purposes (Alvidrez et al., 1996; Okazaki & Sue, 1995). Comprehensive measures have been developed recently for assessing ethnic status. For example, Phinney's (1990, 1996) work provides a model for examining ethnicity as a variable within psychological research. Phinney (1996a) recommends a dimensional model of describing ethnic background that incorporates cultural norms and values, ethnic identity, and minority status as three constructs that represent psychologically meaningful within-group and between-group variation.

Sabnani and Ponterotto's (1992) review of instruments for assessing ethnic minority specific counseling issues is also an informative source of research on ethnic classification. Within the review, the authors evaluate the theoretical basis, development, and psychometric properties of eight ethnic minority specific measures, such as the African Self-Consciousness Scale (ASCS) (Baldwin & Bell, 1985).

A final recommendation for researchers is to report how information on ethnic status was collected and analyzed (Crews & Bindon, 1991). There are numerous techniques for obtaining ethnicity data (e.g., self-report, parent report, hospital chart review). While it is recommended that researchers employ self-report when assessing ethnic status, this strategy presents challenges for child and pediatric studies. Foster and Martinez (1995) point out that it may be difficult to obtain reliable information on ethnicity from younger children. Depending on the descriptive labels being employed, they recommend obtaining ethnic classification from another informant (e.g., parent) for children under the age of 6 years.

Research Design Decisions: Quantitative and Qualitative Methodologies

Traditionally, child clinical and pediatric researchers have relied heavily on quantitative methods and have largely neglected qualitative or emic approaches to study design. While quantitative, or etic, methods offer the statistical and design

benefits of standardization and empirical rigor, qualitative methods generally place more emphasis and focus on the specific belief systems and perspective of the group being studied; consequently, these measures are particularly well-suited for research with ethnic minority populations (Hughes, Seidman, & Williams, 1993). See Chapter 7, this volume, and Krahn, Hohn, and Kime (1995) for a discussion of specific clinical child psychological research studies employing qualitative design.

Recommendations for Research Design and Qualitative or Quantitative Methodology

In order to obtain the benefits of quantitative and qualitative tools in conducting research with ethnic minority children and their families, researchers should incorporate both methodologies whenever this is appropriate to their research question and is feasible (Hines, 1993). Merging methodologies may enable researchers to creatively and sensitively design studies and test hypotheses, particularly when investigating diverse samples (Maton, 1993; Rizzo, Corsaro, & Bates, 1992).

An example of combining qualitative and quantitative techniques in ethnic minority-focused research relevant to children is offered by Bauman and Adair's (1992) description of their study of social support among inner-city mothers of children with a chronic illness. In order to develop a culturally appropriate quantitative measure of social support for this population, the researchers utilized a qualitative phase of ethnographic interviewing. According to the researchers, employing qualitative techniques uncovered previously unanticipated issues that were incorporated into the larger study design. For example, the ethnographic interview data demonstrated that maternal role strain was a major concern for participants, an issue that had been overlooked in the initial study plan. Qualitative analysis of the interview data also guided the wording and language structure that was employed in the questionnaire. According to the researchers, the ethnographic phase enabled them to incorporate the direct experiences of participants, thus substantially enhancing the psychometric sensitivity and strength of future applications of their measure.

Research Design Decisions: Between-Group versus Within-Group Studies

When planning culturally inclusive research, deciding whether a within-group or a between-group design is appropriate involves careful consideration of the benefits and drawbacks to each approach. Within-group designs are studies investigating phenomenon in a specific subgroup, or a single well-defined population. Between-group designs are comparative designs initiated to explore hypotheses concerning the differences or similarities between two or more distinct groups of subjects. Within-group designs allow researchers to gain a detailed assessment of a specific ethnic community; however, they have been criticized for weak designs, a lack of sophisticated hypothesis testing, and for poor generalizability (Hughes et al., 1993). On the contrary, between-group designs are well-suited for statistical analysis and concise hypothesis testing concerning group differences and enable researchers to isolate main effects. However, they may provide minimal insight into the meaning behind identified differences (McKenry et al., 1989).

In the study of ethnic minority groups, psychological researchers have often employed a "race-comparative" design. The race-comparative tradition of comparing minority groups to mainstream samples has been particularly strong within developmental and child clinical literature. For example, between 1936 and 1980, McLoyd and Randolph (1983) found that 47% of all studies involving African-American children in the journal *Child Development* employed a race-comparative design (African-American subjects compared with white children). As opposed to within-group designs which explore individual differences among a specific well-defined population, race-comparative studies investigate hypotheses concerning the differences or similarities between two or more distinct groups of subjects. This research design has been challenged on theoretical and methodological grounds. First, race-comparative designs are based on a tenuous hypothesis that each racially defined group represents a homogeneous sample (LaVeist, 1994). As Zuckerman (1990) has highlighted, more variation may be detected within a single racially defined group than between two different racial groupings. Samples divided on this basis may result in neglect of the considerable range found within ethnic groups [e.g., differences in socioeconomic status (SES), education, acculturation, ethnic identity, age, gender, and setting].

From a theoretical perspective, race-comparative designs limit scientific understanding of psychologically relevant intragroup features (Korchin, 1980; McLoyd & Randolph, 1984). Typically, once a between-group difference is identified, minimal interpretative energy is spent on understanding the meaning and implication of the identified difference (Hughes et al., 1993). Consequently, the prevalence of the race-comparative design has contributed to an inadequately weak knowledge base concerning developmental processes and outcomes within ethnic minority groups.

An additional criticism of between-group, or race-comparative, designs has been the interpretation of "difference" between minority groups and the mainstream as minority group "deficiency" (Korchin, 1980). According to critics, studies using the race-comparative paradigm have reinforced biased accounts of ethnic minorities as deficient along a number of variables in comparison to a white mainstream standard. This deficiency model of highlighting the maladjustment of minority groups has been a major feature of research incorporating diverse populations (Sue, 1993). The focus on ethnicity as an independent variable may create a "person-blame interpretation," inappropriately linking an individual's ethnic background as the primary causal factor in developmental outcome (McLoyd & Randolph, 1984).

Recommendations for Between-Group and Within-Group Study Designs and Models of Successful Research Initiatives

Between-Group Designs. If it is determined that a between-group design is the most applicable for testing their hypotheses, researchers should consider several factors in designing their studies in order to avoid problematic conclusions. First, when comparing two ethnically distinct groups along a particular construct, the variables and measures that assess the construct must be equally relevant for each group (Hughes et al., 1993). Equivalent relevance is defined as whether the constructs or measures hold the same meaning between different ethnic groups and

is important in minimizing the risk of forming biased conclusions (Okazaki & Sue, 1995). If equivalence is not established, group differences should not be interpreted as conclusive and may only be considered tentative or exploratory in nature (Azibo, 1988). Further, substantial pilot work may be needed to provide information concerning the equivalence of variables and measures in the ethnic groups that are studied. Finally, in order to ensure a high-quality between-group design, researchers should seek advice from fellow researchers with expertise working in particular ethnic contexts, or may obtain direct feedback from community leaders or focus groups.

Within-Group Designs. Critiques of the race-comparative tradition have made an important contribution to the field by highlighting the benefits of single-group, noncomparative designs aimed at increasing the knowledge base concerning specific ethnic minority groups (Yee et al., 1993). Moreover, a number of researchers have conducted research using sophisticated within-group designs with pediatric and clinical child populations. For example, Stanton and colleagues' work serves as a model for designing a high-quality single-group research initiative focused on a specific ethnic minority population, in this case, urban African-American adolescents (Li, Fang, Stanton, Feigelman, & Dong, 1996; Stanton & Galbraith, 1994; Stanton, Li, Black, Ricardo, & Galbraith, 1994a). Stanton and colleagues' research has served to expand the limited knowledge base on behavior and decision making in ethnic minority populations at an elevated risk for adverse consequences [e.g., human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS)] from sexual activity (Stanton et al., 1994a,b, 1997). One of Stanton and colleagues' contributions to the field was the development of a culturally sensitive risk behavior assessment questionnaire [the Youth Health Risk Behavioral Inventory (YHRBI)], utilizing anthropological, theoretical, and developmental perspectives (Stanton et al., 1995). The measure was developed as a means of identifying both protective and AIDS-related risk behaviors through qualitative and quantitative techniques, multiple statistical and conceptual methodologies, and community involvement. The instrument development process was initiated in response to recognition of the critical need to evaluate the impact of educational programming on urban African-American youth. Following identification of the need for such a measure, the researchers conducted three research phases: The first phase involved ethnographic data collection that recruited input and guidance from a variety of community members. A coalition of key informants, including child representatives, was formed to help guide the research process. In addition, focus groups were convened to provide preliminary recommendations for the structure, concentration, and format of the measure. The second phase involved designing and refining the actual measure and relied heavily on pilot testing and "debriefing" interviews following administration of the measure to elicit subject's reactions and recommendations. The final phase of measure development involved quantitative assessment of the psychometric properties of the instrument. Throughout the measure development process, Stanton and colleagues (1995) employed multiple methodologies and techniques in order to develop a tool tailored along a culturally inclusive paradigm.

The development of the risk assessment inventory has enabled the research team to conduct comprehensive descriptive studies on sexual practices and risk

behaviors within African-American youth (Stanton et al., 1994a,b). In addition to survey-based studies identifying the sexual behaviors and influences within this cohort of at-risk youth, Stanton and colleagues conducted a longitudinal study examining shifts in risk behaviors within a group of African-American youth over a period of 2 years (Stanton et al., 1997). This valuable research has served to inform educational and intervention programming for ethnic minority children and adolescents.

Neal and colleagues' work investigating anxiety disorders within African-American populations is another model for designing within-group research that is highly relevant to clinical child psychology populations (Neal, Lilly, & Zakis, 1993; Neal, Rich, & Smucker, 1994). By examining the prevalence and manifestation of anxiety disorders specifically within this population, Neal and colleagues have provided a clinically relevant perspective on the diagnosis, epidemiology, course, and outcome of anxiety disorders for African-Americans (Neal & Turner, 1991). From a developmental perspective, Neal and colleagues (1993) have examined fears among African-American children and have incorporated sophisticated between-group design to examine ethnicity and childhood fears. Borrowing from a between-group design, the researchers examined and contrasted the fears reported by African-American children and those reported by white children to determine if clinically relevant differences could be detected. The findings revealed a great deal of similarity between the groups; however, clinically relevant differences in some anxiety profiles were identified for African-American and white children. For example, while a majority of reported childhood fears were similar across groups, differences such as the lack of school fears among African-American children were identified.

Research Implementation Challenges

While the critical need for more child clinical and pediatric psychology research focused on cultural diversity is well established, conducting research with ethnic minority children and families presents numerous methodological challenges. In particular, implementing such research presents a range of difficulties including sampling, subject recruitment, and data collection issues, which are discussed in this next section.

Sampling Issues

Investigators who are interested in conducting ethnically inclusive studies are faced with very difficult sampling decisions. Foster and Martinez (1995) outline two major approaches to sampling in ethnically diverse research, each with their own benefits and drawbacks: convenience and representative sampling strategies. The first possible strategy is to rely on convenience sampling, which does not require a plan for specific ethnic minority participant recruitment. However, researchers who use a convenience sample generally are not able to examine the role of ethnicity in their findings. Further, one risks obtaining a highly nonrepresentative sample, which limits generalizability of findings. One example of convenience sampling within child clinical psychology is utilization of school systems for accessing subjects. Relying on schools for convenience sampling excludes children

with poor attendance or those children who have dropped out of school, which may limit the involvement of certain participants, including ethnic minorities, and interfere with the generalizability of the findings.

The second main sampling alternative is representative sampling. As described by Foster and Martinez (1995), in representative sampling, researchers recruit participants in order to match the demographic composition of the larger population. This procedure allows researchers to obtain a study sample reflective of the general population's ethnic breakdown. This approach helps avoid making inaccurate generalizations concerning diverse groups based on a homogeneous sample. While representative sampling offers greater generalizability of findings than convenience sampling, data will remain most generalizable for the majority population represented in the data collection. Furthermore, representative sampling requires a substantial amount of effort and potential cost in order to identify, recruit, and retain appropriate participants (Sasao & Sue, 1993).

Recommendations for Sampling and Accessing Diverse Populations

When conducting research with ethnically diverse populations, representative sampling is the most appropriate strategy. Despite the fact that representative sampling does not guarantee generalizability of findings, it is a preferable approach, particularly when conducting epidemiological or descriptive research (Foster & Martinez, 1995). Further, in order to secure funding support for research, inclusion of ethnic minorities may require representative sampling to accurately reflect the demographic characteristics of a target population of interest (e.g., children with asthma) (Hohmann & Parron, 1996; National Institutes of Health, 1994).

Subject Recruitment Issues

In order to secure samples that are more representative of ethnically diverse populations, researchers need to understand the special challenges and barriers that are involved in recruiting ethnic minority children and adolescents for research protocols. For example, parents may be hesitant to offer informed consent on behalf of their children if they do not understand or trust the intentions of a research protocol. Recruitment of ethnic minority subjects may be difficult due to the historically tenuous relationship between ethnic minority communities and scientific research. For example, the impact of the Tuskegee syphilis experiment still resonates in public opinion and has tainted the reputation of scientific research protocols among African-Americans (Jenkins & Parron, 1995; Jones, 1993).

Another challenge faced by researchers who are recruiting subjects from ethnic minority and possibly underprivileged communities is parental concern about the risk of being judged as ineffective or even neglectful caretakers. Despite the rules of confidentiality and the nonjudgmental stance of researchers, ethnic minority parents may feel (and be) vulnerable to scrutiny or retribution if they agree to participate in a child-focused study. Apprehension of judgment or pathologization may impede subject recruitment and may limit data collection if subjects do not respond honestly and openly to questions. One example of this problem is research with immigrant families who may fear deportation or disclosure of illegal immigrant status.

An additional example of the considerable logistical challenges associated with subject recruitment and data collection procedures with ethnic minority families involves research with American Indian children and families. The more than 200 American Indian independent tribes within the United States represent enormous linguistic and cultural diversity (Flack et al., 1995). American Indian tribal groups are geographically scattered, and a majority include fewer than 1000 members (Norton & Manson, 1996). In addition, American Indian communities have faced tremendous discrimination and exploitation, resulting in hesitancy to be investigated by researchers. Norton and Manson (1996) reported that conflicts between tribal leaders and researchers have led to widespread mistrust of research. Some tribes have experienced negative consequences from research participation, and tribal representatives have openly questioned the benefits of studies, establishing independent review boards to screen and monitor research protocols (Norton & Manson, 1996).

Recommendations for Subject Recruitment

The most effective means of recruiting an ethnically diverse sample is through careful planning and persistence. To maximize recruitment of ethnically diverse samples, it is best to avoid relying on a single setting for subject recruitment efforts. For example, when conducting ethnically inclusive pediatric psychology studies, it may be necessary to recruit subjects from a range of health care settings. In the planning of one's research, it is important to anticipate the need for multisite recruitment and the additional work associated with such an endeavor (e.g., additional Institutional Review Board submissions, building collaborative relationships with different health care professionals).

To best implement research with ethnically diverse populations, an additional recommendation is to involve consultants in preparing data collection protocols to facilitate cultural sensitivity and acceptance by the community. Gaining input from community leaders, professionals who have experience with cross-cultural psychology and members of other fields such as anthropology and sociology can help prevent problems and serve to promote successful data collection and solid theoretically grounding. Sturm and Gahagan (1999) recommend involving consultants to serve as "interpreters of culture" to provide advice on how to develop and tailor research protocols to specific communities.

When working with ethnic minority populations, it may be wise to use cultural experts at each stage of the research process. Community consultants, or multiple stakeholders, initially should be surveyed to help determine the research needs of underserved populations (Hughes et al., 1993). Enabling community advocates to have a role in establishing research priorities and data collection strategies can help promote participation in and acceptance of research programs (Okazaki & Sue, 1995). Recommendations for research with American Indians, in particular, include the importance of involving tribal representatives and cultural experts in psychological research with this population (LaFromboise, 1988; Norton & Manson, 1996). Additional recommendations and examples of successful research initiatives are included in the *Journal of Consulting and Clinical Psychology* special section entitled "Recruiting and retaining minorities in psychotherapy research" (October 1996).

Data Collection Issues

Language and Cultural Barriers

Once access to a representative diverse sample is established, researchers face additional barriers in implementing ethnically inclusive research designs. First, language barriers may present difficulties. Because of measurement considerations, research designs may require a high level of literacy and English competency as inclusion criteria. In order to administer standardized measures and informed consent procedures without the methodological complications associated with translation and the use of interpreters, researchers often limit their sample to participants with English skills. This results in a biased sampling toward more highly acculturated families (Molina & Chassin, 1996). For example, in research with Spanish-speaking children, researchers may limit their subject pool to families where children and parents both possess English skills, limiting generalizability to the wider community. Difficulties in cross-cultural communication styles and strategies have also been identified as problematic in psychological assessment of African-Americans, particularly in the area of cognitive testing (Miller-Jones, 1989; Helms, 1992).

Issues Related to SES

In addition to linguistic challenges, other logistic factors make it difficult to implement ethnically diverse child clinical and pediatric research initiatives. For example, when targeting a culturally diverse sample within a low SES population, researchers should anticipate difficulty in identifying eligible families and in maintaining contact with research participants. Lower SES families, in general, and families in immediate financial crisis, in particular, may present a range of challenges including the lack of a telephone and homelessness. Conducting longitudinal studies with children from families under stressful and chaotic living situations creates feasibility challenges (Kato & Mann, 1996). Furthermore, the correlation between low SES and ethnic minority status is high in some pediatric populations such as sickle cell disease (Lemanek, Buckloh, Woods, & Butler, 1995).

Measurement Concerns

When conducting cross-cultural research or studies that incorporate an ethnically diverse subject pool in child clinical or pediatric research, a number of difficult psychometric considerations must be addressed to enhance the validity of findings for these populations. The process of exporting measures between cultural settings and applying norms to a range of ethnic groups involves a series of steps to demonstrate cross-cultural equivalence. Okazaki and Sue (1995) highlight the importance of *linguistic*, *conceptual*, and *metric* equivalence in cross-cultural research. While they are often neglected, these concepts are equally important to consider in designing and implementing research with diverse ethnic minority populations of children and adolescents within the United States. First, linguistic equivalence involves ensuring that measures have been translated appropriately, retaining the identical psychometric properties between language versions. Second, conceptual equivalence of measurement relates to whether the psychological

phenomenon or theory possesses the same meaning in each group being evaluated. Finally, the authors recommend demonstrating metric equivalence between groups. Prior to comparing groups based on a standard system of norms, a researcher must ensure that the system holds comparable validity between groups. For example, the statistical properties of a measure (e.g., mean and standard deviation) should be equivalent between groups or statistically adjusted to handle any discrepancies. Hughes and colleagues (1993) also highlight the importance of demonstrating scale equivalence, which requires that response options hold equivalent meaning across groups. Studies have shown that response patterns to Likert-type scales, in particular, may differ between ethnic groups. For example, a stronger tendency to endorse extreme response options has been demonstrated in African-American high school students and Hispanic adolescent/young adults in comparison to white samples (Bachman & O'Malley, 1984; Hui & Triandis, 1989). Hughes and colleagues (1993) also discuss the need for *task equivalence*, which means that respondents are equally familiar and comfortable with the assessment situation. Cultural factors have been shown to play a role in how situational research demands are interpreted, and ethnic background may play an influential role in how children and families respond to different testing demands (i.e., interview formats, self-report questionnaires, naturalistic observations).

The development of measures that are valid for ethnic minority samples of children and families in pediatric and child clinical populations is a tall order, and researchers may wish to follow the lead of Stanton and colleagues (1995) in developing culturally relevant and sensitive measures for children and adolescents. Moreover, Flaherty and colleagues (1988) have provided useful suggestions for establishing cross-cultural equivalence of instruments that are also useful in measurement applications with ethnically diverse samples of children and adolescents in the United States.

When selecting measures for inclusion in a study, in addition to attempting to demonstrate cross-cultural equivalence of traditional psychological measures, it is important to survey a range of available measurement possibilities. Phinney (1990, 1996b), whose work has focused on developing culturally inclusive measures of ethnic identity development, illustrates the importance of establishing cross-cultural equivalence in research with children and adolescents in the United States. Phinney's contributions have included a valuable analysis of the stages of ethnic identity development, the role of ethnic identity in relationships between minorities and the dominant culture, as well an exploration of acculturation in child development (Phinney & Chavira, 1995; Phinney & Onwughalu, 1996; Phinney, Chavira, & Williamson, 1992). Perhaps her most significant work related to child clinical and pediatric populations has been her research on ethnic identity and self-esteem in minority youth (Phinney, 1991, 1996a; Phinney, Chavira, & Tate, 1993). In a longitudinal study incorporating African-American, Asian-American, and Hispanic 16- to 19-year-old subjects, Phinney and Chavira (1992) found a trend of increasing stages of ethnic identity over a 3-year span. A key finding was a significant relationship between self-esteem and ethnic identity throughout the age span studied. Furthermore, Phinney (1992) has developed the Multigroup Ethnic Identity Measure, a scale to measure ethnic identity within diverse populations. The questionnaire was developed using a highly diverse sample, and its psychometric properties were determined based on the scale's sensitivity and appli-

cability to a diverse range of adolescents and adults. Phinney and colleagues' approach to measure development is a productive model for establishing the psychometric properties of an instrument in an ethnically diverse sample.

IMPLEMENTING CROSS-CULTURAL RESEARCH: LESSONS FROM A STUDY OF UGANDAN INFANTS AND THEIR FAMILIES

Pediatric and child clinical psychologists may be called on to apply psychological methods to diverse pediatric populations of varying cultures and ethnic backgrounds, including those outside of the United States. An example of a cross-cultural research initiative that illustrates some of the issues in implementing research to address some of the methodological problems highlighted previously was conducted by one of the present authors (Drotar and colleagues) on the developmental outcomes of Ugandan infants with HIV infection. We present this in some detail in hopes that the lessons learned in the course of this work will be useful to others.

Background on the Impact of HIV Infection on the Development of Ugandan Infants

The need for and scientific significance of this research arose from the following: The large numbers of women and infants who are affected by HIV infection in Uganda, as in other areas of Africa, have assumed epidemic and tragic proportions (Goodgame, 1990). Moreover, the large numbers of children who are affected with HIV infection in Uganda provided a special opportunity to describe and understand the natural history of this condition, including the impact on their cognitive and motor development. It has been well recognized that HIV infection has significant impact on children's cognitive and neurological development (Aylward, Butz, Hutton, Joyner & Vogelhut, 1992; Belman, 1990; Gay et al., 1995). However, the conclusions that can be drawn from studies of the impact of HIV infection on children's development in the United States have been limited by significant methodological problems such as small sample size, a lack of controls, and presence of confounding factors such as prematurity, low birth weight in children, as well as maternal substance abuse. As a consequence, it has been difficult to obtain a clear picture of the natural history of the developmental outcomes associated with HIV infection independent of confounding influences. Such research is possible in Uganda where maternal drug addiction is largely nonexistent.

To address this need, our research group designed and successfully implemented a National Institutes of Health (NIH)-funded study of the cognitive developmental and health outcomes of Ugandan infants. The design for the project included a follow-up of three groups of infants: (1) HIV-infected infants from HIV-infected mothers; (2) non-HIV-infected infants who were born to HIV-infected mothers who are termed *seroreverters*; and (3) a comparison group of uninfected infants born to uninfected mothers. The design called for a prospective follow-up of the cohort of mothers and infants from birth to 24 months using a comprehensive assessment of infant cognitive and motor developmental, neurological and health status. In addition, given our interest in assessing the potential impact of maternal

caretaking influences on children's development, we utilized home observational methods to assess caretaker–infant interaction. The methods and outcomes of the current study are described in detail elsewhere (Drotar et al., 1997).

Measurement and Psychometric Issues

In the course of developing the project, we encountered several problems that are common to applications of psychological methods to other cultures and ethnic minority groups (Berry, Poortinga, Segall, & Dagen, 1992). For example, one problem that we faced was the choice of measures that would be potentially applicable to Ugandan infants. Because we were interested in using objective measures of cognitive and motor development that had adequate standardization, we chose the Bayley Scales (Bayley, 1967). (The study was done prior to the development of the second edition of the Bayley Scales.) We were also interested in extending our assessment of the impact of HIV infection beyond traditional measures of infant sensorimotor development to infants' information-processing skills. To accomplish this goal, we used the Fagan Test of Infant Intelligence (FTII) (Fagan, Singer, Montie, & Shepherd, 1986), because it provided an assessment of infant abilities that did not require any motor responses and had shown adequate predictive validity for the prediction of mental retardation in several samples (Fagan et al., 1986; Fagan & Detterman, 1992; McCall & Carriger, 1992). Despite the potential advantages of the above tests in assessing the development of infants in the United States, we could not be sure that they were applicable to Ugandan infants. Consequently, we needed to conduct feasibility studies to enhance cultural relevance and applicability.

Modifying Procedures to Enhance Cultural Relevance

Linguistic Equivalence

Our research group utilized several procedures to ensure that our measures would be culturally relevant (Berry et al., 1992). First, instructions for the Bayley Scales were translated into Luganda, the language used by the families in Kampala, Uganda, and surrounding areas. Our research group standardized the language of instructions to infants so that they could be applied uniformly for the entire sample. We also decided to employ Ugandan research assistants to administer all tests because they were familiar with the language and culture and would be more acceptable and familiar to the research participants. Fortunately, educated Ugandans all learn to speak English, which facilitated our task of training research assistants to conduct the protocol.

Task Equivalence

We conducted extensive pilot testing of both measures prior to implementing the study protocol on 60 physically healthy Ugandan infants between the ages of 3 and 24 months. This testing revealed several observations: First, the items from the Bayley Motor Scale were readily transferable to Ugandan culture. Ugandan infants had no problems completing these tasks. In fact, they were precocious in doing so, which was consistent with previous research (Evans, 1970). In addition,

most of the items from the Mental Scale that involved early developmental skills, such as visual tracking, reaching, and playing with simple objects, also could be readily completed by Ugandan infants who again were precocious relative to US norms. On the other hand, pilot testing revealed specific problems in applying the Bayley Mental Scale to Ugandan infants that required procedures to be modified. First, the language-related items were not applicable because they included pictures of certain objects such as a car that were not commonly found in Ugandan culture. Because one would not expect Ugandan infants to recognize these items, they would not be sufficiently sensitive. Consequently, we hired a local artist to develop picture cards of objects such as a chair, a goat, and a dress that were familiar to most Ugandan infants. These items were assessed in pilot testing and found to be familiar to infants.

Pilot testing also revealed another interesting observation: several of the more advanced items on the Bayley Mental Scale involved giving infants instructions that required their response to a small rubber doll. For example, infants were asked to point out body parts on the doll, as well as to carry out simple instructions using the doll, such as "put dolly on the chair," "give the dolly a drink." However, Ugandan infants responded quite negatively to the standard doll that was found in the Bayley test kit. Most infants avoided the doll and some even cried when the doll was presented to them. Consequently, if we were going to include items on the Bayley Mental Scale that involved comprehension of language, we needed to come up with an alternative doll that would be more acceptable to the Ugandan infants. Through trial and error, which involved presentation of several different versions, we eventually identified a straw doll that was available in the local villages, was familiar to the infants, and hence acceptable to them.

The next hurdle that we faced was to test whether the FTII was feasible for use in Uganda and would be accepted by Ugandan infants. The FTII includes stimuli measures in which various human faces (infants and adults) are presented in pairs to infants. While the original version of the FTII included Caucasian faces, alternative forms that include individuals of various races have been developed. Because we were concerned that Ugandan infants might have difficulty responding to the stimuli involving the faces of Caucasians, who are novel in Uganda, we utilized several forms of the Fagan test that involved stimuli with individuals of different races. We found that Ugandan infants demonstrated similar visual recognition for novelty scores as infants in the United States and in other cultures (Fagan et al., 1991). This finding not only gave us confidence that we could apply the FTII in Uganda but was interesting and potentially important in itself, in that it underscored the comparable information-processing abilities of infants in different cultures (Fagan et al., 1991).

Metric Equivalence

As is typical in a procedure for using the Bayley Scales, the number of items passed is used to derive a mental development index (MDI) or a physical development index (PDI), which is the variable measure that is usually utilized in many research studies (Bayley, 1969). However, when we summarized the results from the MDIs and PDIs, we found that Ugandan infants tended to score much higher (about one standard deviation) than US infants during the first 15 months of their

lives. Given such differences in test scores, we felt that the MDIs and PDIs based on US norms should not be used as the primary dependent measures in our data analysis. For this reason, we decided to use the number of items that were correctly passed by infants as the primary variable in all data analyses. This allowed us to fulfill our study aims of comparing groups in a way that did not provide misleading information concerning the absolute level of infants' performance.

Identifying Culturally Relevant Norm Groups

We still faced the problem of deciding what norms to use in determining a normal versus an abnormal test performance. We elected to use data from a control group of infants who were born to HIV-infected mothers as a reference group to identify infants who had abnormal test scores, defined as two standard deviations below the mean (Drotar et al., 1997).

Process of Identifying Culturally Sensitive and Culturally Relevant Measures

In developing the methods for our studies of impact of HIV infection in Ugandan infants, our final challenge was to find or develop an assessment of mother–caretaker–infant interaction and caretaking that could be used to as an independent assessment of frequency and quality of caretaking. Such a measure was needed to assure ourselves that whatever findings we obtained concerning the impact of HIV infection on cognitive and motor development could not be explained by differences in the frequency and quality of mother–child interaction and caretaking. This was important because one would expect mothers of HIV-infected infants to have more difficulty providing caretaking and maintaining positive interactions with their children owing to the disruption imposed by their HIV infection. While most mothers were not symptomatic at the time we conducted the study, some had more advanced illness.

In order to assess mother–infant interaction objectively and reliably, we searched for possible measures that could be utilized in a range of home settings including villages where many of the mothers and infants live in Uganda. Sigman and colleagues (1989) utilized such measures in their studies of the health and cognitive development of infants in Kenya. Based on our knowledge of Ugandan mothers and infants, we felt that the measures employed in these studies would be applicable to assessing mother and family caretaker–infant interaction in Uganda. Our pilot data with 30 families with infants ranging in age from 3 to 12 months indicated that indeed the measures employed by Sigman et al. (1989), which included observations of touches, vocalization, and physical care, allowed us to record a range of caretaker–infant interactions that occurred among Ugandan infants, their mothers, and other family caretakers. We also found that this observational method was well accepted by Ugandan families.

In addition to using a measure of the frequency of mother–infant interaction, we also were interested in identifying a measure of the quality of stimulation that was provided by mothers to their infants, which has been consistently linked to positive cognitive development in infants and young children (Elardo & Bradley, 1981). The most widely used measure in the United States of the quality of stimula-

tion provided by caretakers is the Home Observation for the Measurement of Environment Scale (HOME) (Caldwell & Bradley, 1979). The HOME is based on standardization data gathered from children in the United States. However, again, we could be sure that this measure was applicable to infants in Uganda unless we assessed the applicability measure directly. Our pilot testing on 30 infants ranging in age from 3 to 12 months indicated that some of the scales of the measures such as Maternal Responsivity, Involvement, and so forth, which involved degree of maternal responsiveness and affect, were quite applicable to Ugandan mothers and infants. On the other hand, other scales such as Provision of Play Materials were not applicable because they required assessment of objects in the home, such as books, that were not readily available in Uganda. Consequently, we used only those scales that were judged to be most applicable to the assessment of the quality of the home environment for infants in Uganda.

Importance of Extensive Pilot Testing in Cross-Cultural or Ethnically Diverse Research Initiatives

Our experience in studies of the development of Ugandan infants has underscored the importance of conducting detailed pilot work when applying measures to various groups or individuals who were not included in the initial standardization sample of an instrument. While the need for such studies is obligatory in research in other countries, researchers in pediatric and clinical psychology also encounter important variations in culture and ethnicity in research with populations of children in the United States for whom validity and acceptability of measures should not be automatically assumed and needs to be established empirically.

IMPLICATIONS FOR TRAINING AND PROFESSIONAL DEVELOPMENT

The present chapter has clearly underscored the need for an expanded focus on ethnic diversity within child clinical and pediatric psychology research. Increasing ethnic diversity in research with pediatric and child clinical populations will provide the opportunity to develop more successful intervention and prevention efforts and will enable psychologists to serve as culturally informed child advocates. Despite the demonstrated need for an increased concentration on ethnic diversity within psychology, multiculturalism has been largely neglected in graduate training. Evaluations of the curricula of accredited clinical psychology training programs have revealed insufficient coverage of ethnic diversity issues (Bernal & Castro, 1994; Bernal & Padilla, 1982). In addition, the profession of psychology remains largely homogeneous according to surveys of the American Psychological Association's membership. Approximately 4% of the American Psychological Association is composed of ethnic minority professionals, which is not representative of the general population's demographic composition. Overall, the field of psychology has failed to adequately address issues of ethnicity, resulting in a need for more training and professional development opportunities tailored to research or clinical intervention with diverse populations (Bernal & Castro, 1994; Hall, 1997). As the United States grows increasingly diverse, the need for multiculturally competent psychologists will become even more pressing. Furthermore, cross-

cultural training for mental health professionals has been shown to enhance the clinical skills, service delivery, treatment outcome, and cost-effectiveness of trainees (Lefley, 1984).

The need for training in multicultural issues is equally critical for students within child clinical and pediatric psychology (Vargas & Willis, 1994). Children and adolescents will continue to comprise the most ethnically diverse segment of the US population. Consequently, students trained in child clinical and pediatric psychology need to be prepared to conduct research that serves to improve our knowledge base concerning ethnic minority children and families. The following section summarizes recommendations for meeting this formidable yet critical educational challenge.

Increasing Professional Diversity and Development

An important preliminary step in responding to the need for an increased focus on the research and clinical needs of ethnic minority children and families is to assess the level of diversity within the field itself. The relative homogeneity found within the field mandates the need to attract and support young ethnic minority investigators and academicians (Bernal & Castro, 1994; Betancourt & Lopez, 1993; Hall, 1997). This mandate will require concerted efforts at identifying talented minority students very early in their careers and encouraging them to pursue careers in child clinical and pediatric psychology, ideally with the support of stipends for research training. In addition to facilitating ethnic minority faculty recruitment, an important recommendation for increasing professional development is to provide continuing education programming related to research and intervention with ethnically diverse populations for practicing psychologists (Ricardo & Holden, 1994).

Promoting a Focus on Ethnic Minority Issues in Child Clinical and Pediatric Psychology Research and Clinical Training

In their introduction to a special section of the *Journal of Clinical Child Psychology* devoted to treatment and assessment issues for ethnic minority adolescents, Vargas and Willis (1994) described difficulty in soliciting an adequate number of submissions. This anecdote demonstrates the need for increased focus in clinical and research training concerning adolescents. This can be accomplished only by stimulating and supporting a specific focus on research and clinical training concerning culture and ethnicity in child clinical and pediatric psychology. Specific recommendations for accomplishing this task include curriculum development, mentoring initiatives, and fostering interdisciplinary exposure. Program directors have an opportunity to shape the field by incorporating cultural issues into existing syllabi and by developing courses (e.g., research methods and assessment, assessment and intervention, etc.) with a focus on ethnic minority issues. Further, clinical and research training programs should offer students material exposure to ethnically diverse child clinical and pediatric populations in practicum, internship, and postdoctoral experiences. In addition, psychology departments should offer extensive mentoring support to ethnic minority trainees and to students with special interests in cultural issues. See Atkinson, Casas, and

Neville (1994) and McNeill, Hom, and Perez (1995) for a discussion of mentoring. Finally, students should have the opportunity to become exposed to culturally relevant interdisciplinary training experiences. For example, collaborating with faculty and students from other disciplines such as anthropology, sociology, epidemiology, and public health is particularly beneficial when conducting child clinical and pediatric research with ethnically diverse populations.

Developing Model Training Programs That Focus on Research with Ethnic Minority Populations

Another recommendation to advance research training is to develop model programs that are focused on the special research needs of ethnic minority populations of children and adolescents. Such programs are few and far between. One example is the University of Miami's (school of medicine and the department of clinical psychology) training program, "Health Behavior Research in Minority Pediatric Populations." The purpose of the program is to emphasize the role of psychosocial factors in health conditions impacting ethnic minority youth. The proposed curriculum includes an interdisciplinary focus to address the complex set of research issues and needs facing ethnic minority children and families. A primary advantage of this program is the availability of diverse populations of ethnic minorities in this setting, which maximizes clinical and research opportunities with these populations. The University of Miami's research training program serves as a model of responding to the need for an increased focus on ethnic and cultural diversity by developing specialized training opportunities that should stimulate the career development and expertise of researchers to work with children and adolescents from diverse cultural and ethnic groups. We encourage other programs to follow their lead.

REFERENCES

- Alvidrez, J., Azocar, F., & Miranda, J. (1996). Demystifying the concept of ethnicity for psychotherapy research. *Journal of Consulting and Clinical Psychology*, 64, 903-909.
- Atkinson, D.R., Casas, A. & Neville, H. (1994). Ethnic minority psychologists: Whom they mentor and benefits they derive from the process. *Journal of Multicultural Counseling and Development*, 22, 37-48.
- Aylward, E. H., Butz, A. M., Hutton, N., Joyner, M. D., & Vogelhut, J. W. (1992). Cognitive and motor development in infants at risk for human immunodeficiency virus. *American Journal of Diseases of Children*, 146, 218-222.
- Azibo, D. (1988). Understanding the proper and improper usage of the comparative research framework. *Journal of Black Psychology*, 15, 81-91.
- Bachman, J. G. & O'Malley, P. M. (1984). Yea-saying, nay-saying, and going to extremes: Black-white differences in response styles. *Public Opinion Quarterly*, 48, 491-509.
- Baldwin, J. A., & Bell, Y. R. (1985). The African self-consciousness Scale: An Africentric personality questionnaire. *Western Journal of Black Studies*, 9, 61-68.
- Bauman, L. J., & Adair, E. G. (1992). The use of ethnographic interviewing to inform questionnaire construction. *Health Education Quarterly*, 19, 9-23.
- Bayley, N. (1969). *Bayley Scale of Mental Development*. New York: Psychological Corp.
- Belman, A. I. (1990). AIDS and pediatric neurology. *Neurologic Clinics*, 6, 571-603.
- Bernal, M. E., & Castro, F. G. (1994). Are clinical psychologists prepared for service and research with ethnic minorities? Report of a decade of progress. *American Psychologist*, 49, 797-805.

- Bernal, M. E., & Padilla, A. M. (1982). Status of minority curricula and training in clinical psychology. *American Psychologist*, 37, 780–787.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dagen, P. R. (1992). *Cross-cultural Psychology. Research and Applications*. New York: Cambridge University Press.
- Betancourt, H., & Lopez, S. R. (1993). The study of culture, ethnicity, and race in American psychology. *American Psychologist*, 48, 629–637.
- Beutler, L. E., Brown, M. T., Crothers, L., Booker, K., & Seabrook, M. K. (1996). The Dilemma of Factitious Demographic Distinctions in Psychological Research. *Journal of Consulting and Clinical Psychology*, 64, 892–903.
- Caldwell, B. M., & Bradley, R. H. (1979). *Home observation for measurement of the environment*. Little Rock: University of Arkansas.
- Carter, R. T. (1995). *The influence of race and racial identity in psychotherapy: Toward a racially inclusive model*. New York: John Wiley.
- Crews, D. E., & Bindon, J. R. (1991). Ethnicity as a taxonomic tool in biomedical and biosocial research. *Ethnicity and Disease*, 1, 42–49.
- Drotar, D., Olness, K., Wiznitzer, M., Guay, L., Marum, L., Svilar, M. S., Hom, D., Fagan, J. F., Ndugwa, C., & Kiziri-Mayengo, R. (1997). Neurodevelopmental outcomes of Ugandan infants with human immunodeficiency virus type I infection. *Pediatrics*, 100, E1–E5.
- Elardo, R., & Bradley, R. H. (1981). The Home Observation for the Measurement of the Environment (HOME) Scale: A review of research. *Developmental Review*, 1, 113–145.
- Evans, J. L. (1970). *Children in Africa. A review of psychological research*. New York: Teacher's College Press.
- Fagan, J. F., & Detterman, D. (1992). The Fagan test of infant intelligence: A technical summary. *Journal of Applied Developmental Psychology*, 13, 173–193.
- Fagan, J. F., Drotar, D., Olness, K., Peterson, N., Kiziri-Mayengo, R., Guay, L., Ndugwa, C., & Berkoff, K. (1991). *The Fagan test of infant intelligence: Cross-cultural and racial differences*. Paper presented at annual meeting of Society for Behavioral Pediatrics, Baltimore, MD, 1991.
- Fagan, J. F., Singer, L. T., Montie, J., & Shepherd, P. A. (1986). Selective screening device for the early detection of normal or delayed cognitive development in infants at risk for later mental retardation. *Pediatrics*, 78, 1021–1026.
- Flack, J. M., Amaro, H., Jenkins, W., Kunitz, S., Levy, J., Mixon, M., & Yu, E. (1995). Panel I: Epidemiology of minority health. *Health Psychology*, 14, 592–601.
- Flaherty, J. A., Gaviria, G.M., Pathak, D., Mitchell, R., Wintrob, R., Richman, J. A., & Birz, S. (1988). Developing instruments for cross-cultural psychiatric research. *Journal of Nervous and Mental Disease*, 176, 257–263.
- Foster, S. L., & Martinez, C. R., Jr. (1995). Ethnicity: Conceptual and methodological issues in child clinical research. *Journal of Clinical Child Psychology*, 24, 214–226.
- Gay, C. L., Armstrong, D., Cohen, D., Lai, S., Hardy, M. D., Swales, T. P., Morrow, C. J., & Scott, G. B. (1995). The effects of HIV on cognitive and motor development in children born to HIV seropositive women with no reported drug use: Birth to 24 months. *Pediatrics*, 95, 1078–1082.
- Goodgame, R. W. (1990). AIDS in Uganda: Clinical and social features. *New England Journal of Medicine*, 323, 383–389.
- Graham, S. (1992). "Most of the subjects were white and middle class": Trends in published research on African American in selected APA journals, 1970–1989. *American Psychologist*, 47, 629–639.
- Hall, I. C. (1997). Cultural malpractice: The growing obsolescence of psychology with the changing US population. *American Psychologist*, 52, 642–651.
- Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist*, 47, 1083–1101.
- Hohmann, A. A., & Parron, D. L. (1996). How the new NIH guidelines on inclusion of women and minorities apply: Efficacy trials, effectiveness trials, and validity. *Journal of Consulting and Clinical Psychology*, 64, 851–855.
- Hines, A. M. (1993). Linking qualitative and quantitative methods in cross-cultural survey research: Techniques from cognitive science. *American Journal of Community Psychology*, 21, 729–746.
- Hughes D., Seidman, E., & Williams, N. (1993). Cultural phenomena and the research enterprise: Toward a culturally anchored methodology. *American Journal of Community Psychology*, 21, 687–703.
- Hui, H. C., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, 20, 296–309.

- Iwamasa, G. Y., & Smith, S. K. (1996). Ethnic diversity in behavioral psychology. *Behavior Modification, 20*, 45–49.
- Jenkins, R. R., & Parron, D. (1995). Guidelines for adolescent health research: Issues of race and class. *Journal of Adolescent Health, 17*, 314–323.
- Jones, J. H. (1993). *Bad blood. The Tuskegee syphilis experiment*. New York: Maxwell McMillan International.
- Kato, P. M., & Mann, T. (Eds.). (1996). *Handbook of diversity issues in health psychology*. New York: Plenum Press.
- Kawaga-Singer, M., Katz, P. A., Taylor, D. A., & Vanderryn, J. H. M. (1996). *Health issues for minority adolescents*. Lincoln: University of Nebraska Press.
- Korchin, S. J. (1980). Clinical psychology and minority problems. *American Psychologist, 35*, 262–269.
- Krahn, G. L., Hohn, M. F., & Kime, C. (1995). Incorporating qualitative approaches into clinical child psychology research. *Journal of Clinical Child Psychology, 24*, 204–213.
- LaFromboise, T. D. (1988). American Indian mental health policy. *American Psychologist, 43*, 388–397.
- LaVeist, T. A. (1994). Beyond dummy variables and sample selection: What health services researchers ought to know about race as a variable. *Health Services Research, 29*, 1–16.
- Lefley, H. P. (1984). Cross-cultural training for mental health professionals: Effects on the delivery of services. *Hospital and Community Psychiatry, 35*, 1227–1229.
- Lemanek, K. L., Buckloh, L. M., Woods, G., & Butler, R. (1995). Diseases of the circulatory system: Sickle cell disease and hemophilia. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 286–309). New York: Guilford Press.
- Lewis, H. A., & Kliewer, W. (1996). Hope, coping, and adjustment among children with sickle cell disease: Tests of mediator and moderator models. *Journal of Pediatric Psychology, 21*, 25–41.
- Li, X., Fang, X., Stanton, B., Feigelman, S., & Dong, Q. (1996). The rate and pattern of alcohol consumption among Chinese adolescents. *Journal of Adolescent Health, 19*, 353–361.
- Maton, K. I. (1993). A bridge between cultures: Linked ethnographic-empirical methodology for culture anchored research. *American Journal of Community Psychology, 21*, 747–773.
- McCall, R. B., & Carriger, M. S. (1992). A meta-analysis of infant habituation and recognition performance as a predictor of later IQ. *Child Development, 64*, 57–59.
- McKenry, P. C., Everett, J. E., Ramseur, H. P., & Carter, C. J. (1989). Research on black adolescents: A legacy of cultural bias. *Journal of Adolescent Research, 4*, 254–264.
- McLoyd, V. C., & Randolph, S. M. (1983). *Secular trends in the study of Afro-American children: A review of child development 1936–1980*. Biennial Meeting Society for Research in Child Development, Detroit, 1983.
- McLoyd, V. C., & Randolph, S. M. (1984). The content and publication of research on Afro-American children: A content analysis. *Human Development, 27*, 65–75.
- McNeill, B. W., Hom, K. L., & Perez, J. A. (1995). The training and supervisory needs of racial and ethnic minority students. *Journal of Multicultural Counseling and Development, 23*, 246–258.
- Miller-Jones, D. (1989). Culture and testing. *American Psychologist, 44*, 360–366.
- Molina, B. S. G., & Chassin, L. (1996). The parent-adolescent relationship at puberty: Hispanic ethnicity and parent alcoholism as moderators. *Developmental Psychology, 32*, 675–686.
- National Institutes of Health. (1994). *Guidelines on the inclusion of women and minorities as subjects in clinical research*. 59, Fed. Reg. 14,508 (Document no. 94-5435). Washington, DC: US Government Printing Office.
- National Institutes of Health. (1995). *Standards for the classification of federal data on race and ethnicity*. 60, Fed. Reg. 44,673 (Document no. 95-20787). Washington, DC: US Government Printing Office.
- Neal, A. M., Lilly, R. S., & Zakis, S. (1993). What are African American children afraid of? A preliminary study. *Journal of Anxiety Disorders, 7*, 129–139.
- Neal, A. M., Rich, L. N., & Smucker, W. D. (1994). The presence of panic disorder among African American hypertensives: A pilot study. *Journal of Black Psychology, 20*, 29–35.
- Neal, A. M., & Turner, S. M. (1991). Anxiety disorders research with African Americans: Current status. *Psychological Bulletin, 109*, 400–410.
- Norton, I. M., & Manson, S. M. (1996). Research in American Indian and Alaska Native communities: Navigating the cultural universe of values and process. *Journal of Consulting and Clinical Psychology, 64*, 856–860.
- Ogbu, J. U. (1985). A cultural ecology of competence among inner-city blacks. In M. B. Spencer, G. K. Brookins, & W. R. Allen (Eds.), *Beginnings: The social and affective development of black children* (pp. 45–66). Hillsdale NJ: Erlbaum.

- Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment, 7*, 367–375.
- Phinney, J. S. (1990). Ethnic identity in adolescents and adults: Review of research. *Psychological Bulletin, 108*, 499–514.
- Phinney, J. S. (1991). Ethnic identity and self-esteem: A review and integration. Special issue: Ethnic identity and psychological adaptation. *Hispanic Journal of Behavioral Sciences, 13*, 193–208.
- Phinney, J. S. (1992). The Multigroup Ethnic Identity Measure: A new scale for use with diverse groups. *Journal of Adolescent Research, 7*, 156–176.
- Phinney, J. S. (1996a). Understanding ethnic diversity: The role of ethnic identity. *American Behavioral Scientist, 40*, 143–152.
- Phinney, J. S. (1996b). When we talk about American ethnic groups, what do we mean? *American Psychologist, 51*, 918–927.
- Phinney, J. S., & Chavira, V. (1992). Ethnic identity and self-esteem: An exploratory longitudinal study. *Journal of Adolescence, 15*, 271–281.
- Phinney, J. S., & Chavira, V. (1995). Parental ethnic socialization and adolescent coping with problems related to ethnicity. *Journal of Research on Adolescence, 5*, 31–53.
- Phinney, J. S., Chavira, V., & Tate, J. D. (1993). The effect of ethnic threat on ethnic self-concept and own-group ratings. *Journal of Social Psychology, 133*, 469–478.
- Phinney, J. S., Chavira, V., & Williamson, L. (1992). Acculturation attitudes and self-esteem among high school and college students. *Youth and Society, 23*, 299–312.
- Phinney, J. S., & Onwughalu, M. (1996). Racial identity and perception of American ideals among African American and African students in the United States. *International Journal of Intercultural Relations, 20*, 127–140.
- Ponterotto, J. (1988). Racial/ethnic minority research in the *Journal of Counseling Psychology*: A content analysis and methodological critique. *Journal of Counseling Psychology, 35*, 410–418.
- Ponterotto, M., & Casas, J. (1987). In search of multicultural competence within counselor education programs. *Journal of Counseling and Development, 65*, 430–434.
- Reid, P. T., & Kelly, E. (1994). Research on women of color: From ignorance to awareness. *Psychology of Women Quarterly, 18*, 477–486.
- Ricardo, I. B., & Holden, E. W. (1994). Multicultural training in pediatric and clinical child psychology predoctoral internship programs. *Journal of Clinical Child Psychology, 23*, 32–38.
- Rizzo, T. A., Corsaro, W. A., & Bates, J. E. (1992). Ethnographic methods and interpretive analysis: Expanding the methodological options of psychologists. *Developmental Review, 12*, 101–123.
- Rogers-Wiese, M. R. (1992). Racial/ethnic minority research in school psychology. *Psychology in the Schools, 29*, 267–272.
- Rogoff, B., & Morelli, G. (1994) Cross-cultural perspectives on children's development. In P. K. Bock (Ed.), *Psychological anthropology* (pp. 231–242). Westport, CT: Praeger Publishers.
- Sabnani, H. B., & Ponterotto, J. G. (1992). Racial/ethnic minority-specific instrumentation in counseling research: A review, critique, and recommendations. *Measurement and Evaluation in Counseling and Development, 24*, 161–187.
- Sasao, T., & Sue, S. (1993). Toward a culturally anchored ecological framework of research in ethnic-cultural communities. *American Journal of Community Psychology, 21*, 705–727.
- Sigman, M., Neumann, L., Carter, E., Cattle, D. J. D., Souza, S., & Biwilbo, N. (1989). Home interactions and the development of Embu toddlers in Kenya. *Child Development, 59*, 1251–1261.
- Sokol-Katz, J. A., & Ulbrich, P. M. (1992). Family structure and adolescent risk-taking behavior: A comparison of Mexican, Cuban, and Puerto Rican Americans. *International Journal of Addictions, 27*, 1197–1209.
- Stanton, B., Black, M., Feigelman, S., Ricardo, I., Galbraith, J., Li, X., Kaljee, L., Keane, V., & Nesbitt, R. (1995). Development of a culturally, theoretically and developmentally based survey instrument for assessing risk behaviors among African-American early adolescents living in urban low-income neighborhoods. *AIDS Education and Prevention, 7*, 160–177.
- Stanton, B., Fang, X. Li, X., Feigelman, S., Galbraith, J., & Ricardo, I. (1997). Evolution of risk behaviors over two years among a cohort of urban African American adolescents. *Archives of Pediatric and Adolescent Medicine, 151*, 398–406.
- Stanton, B., & Galbraith, J. (1994). Drug trafficking among African-American early adolescents: Prevalence, consequences, and associated behaviors and beliefs. *Pediatrics, 93*, 1039–1043.
- Stanton, B., Li, X., Black, M., Ricardo, I., & Galbraith, J. (1994a). Anal intercourse among preadolescent and early adolescent low-income urban African-Americans. *Archives of Pediatric and Adolescent Medicine, 148*, 1201–1204.

- Stanton, B., Li, X., Black, M., Ricardo, I., Galbraith, J., Kaljee, L., & Feigelman, S. (1994b). Sexual practices and intentions among preadolescent and early adolescent low-income urban African Americans. *Pediatrics*, 93, 966–973.
- Sturm, L., & Gahagan, S. (1999). Cultural issues in provider-parent relationships. In D. Kessler & P. Dawson (Eds.), *Failure to thrive and pediatric undernutrition: A transdisciplinary approach* (pp. 351–374). Baltimore: Paul H. Brookes Publishing.
- Sue, D. W. (1993). Confronting ourselves: The white and racial/ethnic-minority researcher. *Counseling Psychologist*, 21, 244–249.
- Triandis, H. C. (1996). The psychological measurement of cultural syndromes. *American Psychologist*, 51, 407–415.
- Triandis, H. C., & Brislin, R. W. (1984). Cross-cultural psychology. *American Psychologist*, 39, 1006–1016.
- US Bureau of the Census. (1995). Population division, release PPL-57, United States population estimates, by age, sex, race, and hispanic origin, 1990 to 1996.
- US Bureau of the Census. (1995). Current population reports, Series P23–189, *Population profile of the United States: 1995*. Washington, DC: US Government Printing Office.
- Vargas, L. A., & Willis, D. J. (1994). Introduction to the special section: New directions in the treatment and assessment of ethnic minority children and adolescents. *Journal of Clinical Child Psychology*, 23, 2–4.
- Wyatt, G. E. (1991). Examining ethnicity versus race in AIDS related sex research. *Social Science Medicine*, 33, 37–45.
- Yee, A. H., Fairchild, H. H., Weizmann, F., & Wyatt, G. E. (1993). Addressing psychology's problems with race. *American Psychologist*, 48, 1132–1140.
- Zuckerman, M. (1990). Some dubious premises in research and theory on racial differences: Scientific, social, and ethical issues. *American Psychologist*, 45, 1297–1303.

III

Securing Resources for Research

In developing and contributing to this handbook, I decided to place primary emphasis on information concerning knowledge and skills that are necessary for researchers to conduct research in pediatric and clinical child psychology. One of these skills is the ability to secure funding for research. The conduct of clinically relevant research in pediatric and clinical psychology often involves timely and costly procedures involving recruitment and retention of research participants, data collection, and analysis. Moreover, the prospective and multisite research projects that are necessary to advance science in these fields are very expensive. Also, many researchers are under increased pressure to support substantial proportions of their salaries as well as to develop and sustain their research activities through grants. For example, in many applied settings it is very difficult for researchers to obtain sufficient time to conduct research unless they have sufficient research funding to support these activities. Over and above the pragmatic advantages of research funding, a track record of funding is often critical to researchers' career advancement and promotion in university settings.

Although the need for psychologists to conduct research with children and families to develop skills in grant writing is a salient one, grant writing requires a range of skills that are not easily mastered and are not taught in great depth in most graduate training programs. To address this need, Section III includes two chapters that are based on my own experiences in securing funding for my research that focus on helping researchers to understand the process of applications to funding agencies and to develop strategies to maximize their opportunities to obtain such funding. In Chapter 9, readers are given a step-by-step introduction to the critical features of a successful application to foundations, while Chapter 10 presents information concerning preparing grants to secure research funding from government agencies. Each of these chapters include concrete guidelines for researchers who are making applications for research funding, including methods to locate funding sources, communicating with agency and/or foundation staff, and a step-by-step approach in preparing applications for funders. I am hopeful that the detailed description of the application process and what reviewers look for in a research proposal will be useful to readers. Suggestions for training students to develop grant proposals that have been utilized in our graduate training program are also described.

9

Writing Research Proposals for Foundations

DENNIS DROTAR

Foundations are an important source of support for research as well as development of services for clinical child and pediatric populations. For the most part, the application process to foundations is much less involved and detailed than it is for government research grants. This is a clear advantage for a researchers. Nevertheless, several features of foundations make them less accessible to pediatric and clinical child psychologists than they need to be. Because there are such a large number of local and national foundations, it can be difficult for investigators to find out about foundations' interests and priorities. Moreover, each foundation has somewhat different guidelines for the priority areas of the research they fund, as well as for the application and review process. The heterogeneous nature of foundations' structures, procedures, and priorities also can make the application process a confusing one.

Despite the above difficulties, given the level of competition for research funding via government agencies and the availability of foundation support, it is increasingly important that investigators in pediatric and clinical child psychology seriously consider all options that are available to fund their research, including foundations. My review of authors' credits of sources of funding for the 183 research studies that were published in the *Journal of Clinical Child Psychology* and the *Journal of Pediatric Psychology* in 1996 and 1997 indicated that 39 of these studies were supported by foundations. Although these data indicate that some researchers in the above fields have been successful in obtaining research funding from foundations, the level of foundation funding could be even greater.

One salient barrier to obtaining such funding is that many researchers in pediatric and clinical child psychology are not familiar with the process of apply-

DENNIS DROTAR • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

ing to foundations for research funding. Although there are some common elements to writing proposals for government agencies and foundations (Carlson, 1995), the nature of proposals and the application process to foundations are sufficiently different from that required for government grants to warrant special consideration. Moreover, in my experience the level of training and experience received by most psychologists concerning applications to foundations to obtain research funding is minimal. To address this need, the purpose of this chapter is to provide an introduction to the process of developing applications to foundations to obtain research funding and describe practical strategies to facilitate such applications. Interested readers also may wish to consult Carlson (1995), Grant and Bowe (1995), or Fine (1996) for useful information concerning applications to foundations, as well as McIlroy (1998) for an informative description of how foundations operate.

TYPES OF FOUNDATIONS

Carlson (1995) distinguished among four different types of foundations: (1) independent, (2) community, (3) operating, and (4) company sponsored or corporate. Independent foundations are established by individuals or families and include large multipurpose foundations, some of which have extensive staff who are available to review incoming proposals and make recommendations to the foundation's board of trustees. The William T. Grant Foundation, which is based in New York City, is an example of a large, multipurpose foundation that focuses on research concerning children's behavior and mental health. Other types of independent foundations including special purpose foundations that focus on very specific problems and small family foundations that service the needs of their local areas. These latter foundations may have a very small staff, which may include the family members who established the foundation.

In addition to independent foundations, other types of foundations include community foundations that represent the interests and resources of a large number of individuals and groups rather than one family, operating foundations that limit funding within a narrowly defined geographic area such as a city, and corporate foundations, which are established and implemented by a corporation. Some corporate foundations have established sets of priorities in certain areas, for example, early childhood education, that may fit some of the research that is conducted by pediatric and clinical child psychologists.

FINDING OUT ABOUT FOUNDATIONS' RESEARCH PRIORITIES

Table 1 lists steps for researchers in applying to foundations for support. The first set of tasks that are faced by investigators involve locating those foundations that have an interest in funding research and determining whether the specific focus of their research fits the foundation's interests. Carlson (1995) recommends several steps to determine whether there is a suitable match between one's research and foundation priorities. These are: (1) locate foundation funding directories in your library [potential sources include Baumgartner, Feczkó, Hall, & Kovacs (1995);

Eckstein (1996); Fine (1996); Freeman (1991); Garonzik (1985), the *Foundation Directory* (on-line at <http://fdncenter.org>), and the National Network of Grantmakers (1997b)]; (2) identify the subject areas that your research project would fit; (3) use the subject index of the directory to look up the subject area of your project along with foundations in your geographic areas; (4) check the type of support index for the category of support being sought; (5) cross reference the subject index and type of support index in order to identify possible funders; (6) identify each foundation identified in the subject area of your proposal to determine the content areas of foundation interest, as well as type and amounts of potential support; (7) identify those funding sources that best match your project's funding needs; (8) call or write for copies of annual reports, which often contain lists of funded projects; and (9) carefully review the annual reports to identify similarities between the proposed research and research that has been recently funded by foundations.

Using the above procedures, investigators will generally be able to rule out foundations whose priorities clearly do not fit their research and identify potential funders whose priorities may fit their work. In situations where investigators are not entirely clear about the potential fit between a foundation's interest and their research but where it looks possible, a call to the foundation contact person and a letter or proposal describing the research may be very useful to determine whether there is a suitable fit. Whether a call or letter is more appropriate for this purpose will depend on the foundation's application procedures. Foundations' explicit guidelines concerning their preferred mode of contact with investigators should be carefully followed.

Using Multiple Methods to Identify Foundations That Fund Research

Any number of databases concerning foundation support are available, which are very helpful but also can make the search for a potential funding source time consuming (Carlson, 1995; Eckstein, 1996; Freeman, 1991; Garonzik, 1985). Moreover, variations in the clarity of the stated funding priorities included in foundation directories also can make the task of locating potential foundations a more difficult one. Investigators should recognize that in contrast to government agencies that fund research, that is, National Institutes of Health (NIH), many foundations have a diverse set of funding priorities, which may include service development, public education, building and equipment, as well as research. Although the range of projects and content areas that are funded by foundations can sometimes be advantageous to investigators (Fine, 1996), it also can be confusing.

Given potential ambiguities in the listings of foundations' interests and the variation in individual foundations' modes of operation, investigators should utilize multiple methods to target relevant foundations. In my own experience, I have found the statements that foundations publish in their annual reports concerning their priorities and the types of research that they have funded in the past to be particularly helpful. Most foundations publish an annual report in which their goals and priorities for the year are listed along with titles, amounts, and descriptions of funded research. Researchers should carefully study these annual reports to identify specific areas of research have been funded as well as the range of funds that have been awarded to individual investigators.

Table 1. Steps in Applying to Foundations for Research Support

-
1. Identify potential funding sources
 - a. Locate foundation directories in the library and Internet.
 - b. Identify subject areas that the proposed research project fits into.
 - c. Use the subject index of the directory to identify the subject area and foundations in the relevant geographic areas.
 - d. Check the type of support index for the category of support (e.g., research, service) that is relevant to the proposal.
 - e. Cross-reference the subject index and type of support index to identify possible funders.
 - f. Identify foundations in the subject area of the proposal to determine content areas of interest and types and amounts of support.
 - g. Identify those funding sources that have a suitable match to the project's funding needs.
 - h. Call or write for copies of annual reports for foundations.
 - i. Review the foundations' annual reports to determine potential fit between proposed research and foundation priorities.
 2. Before applying for research funds, be sure there is a fit between the content of the proposed research project and the foundation's programmatic interests
 - a. If necessary, contact foundation staff to clarify the fit between the proposed research and foundation priorities.
 - b. Work with Institutional Development Office as relevant.
 3. Understand the specific procedures of foundation review
 - a. Carefully review application forms and guidelines.
 - b. Contact foundation staff to answer questions about the application procedure.
 4. If it is relevant, compose a letter of intent that includes the following:
 - a. Need for the proposed research.
 - b. Experiences and track record of the applicant and organization.
 - c. Research plan.
 - d. Anticipated impact, benefits, or significance of the proposed program.
 - e. Proposed budget.
 5. Develop formal application
 - a. Describe the professional background of the applicant and organization.
 - 1) Describe relevant programs and activities of the applicant's organization, history, and accomplishments.
 - 2) How did the proposed project evolve?
 - b. Describe qualifications of the investigator to carry out the proposed research.
 - 1) What special skills and experience of the investigator and research team will facilitate the proposed research?
 - 2) Has the researcher successfully conducted similar research?
 - c. Describe the need for the proposed research.
 - 1) How many children and families are affected by the problem that will be studied?
 - 2) What are the costs to society of the problem?
 - 3) How will the proposed research address this problem?
 - d. In the event the proposed project is successful, how will the information be used to help children and families?
 - e. Will the project have significance beyond the local community? If so, how?
 - f. Identify the goals and the products of the proposed research.
 - 1) What specific knowledge will be gained from the study?
 - 2) How will the knowledge gained be used to benefit children in families in the local community and society?
 - g. Description of methods.
 - 1) How will the research be conducted?
 - 2) What will be done to enhance the feasibility of the project?
 - 3) What specific methods (e.g., measures, interviews) will be used to conduct the project?
-

Table 1. (*Continued*)

-
- h. Evaluation of proposed research.
 - 1) What outcomes will be evaluated?
 - 2) Why are these outcomes important?
 - 3) How will the evaluation be conducted?
 - 4) What specific benchmarks will be used to determine whether the project will be successful?
 - i. Budget.
 - 1) What funds are necessary to conduct the proposed research?
 - 2) How do specific proposed budget items relate to the specific activities that are proposed?
 - 3) How will the project be funded when foundation support is no longer in place?
 - j. Preparing response to feedback in letter and/or site visit.
 - 1) Answer questions clearly and thoroughly.
 - 2) Supply all information that is requested.
-

Investigators also should consult published articles concerning their area of research interest to identify foundations that are potential sources of funding. It is also extremely helpful for investigators to identify colleagues in their institution or elsewhere who have had experience with and/or received funding from specific foundations and request their consultation concerning proposal review and the application process.

MAKING CONTACT WITH FOUNDATIONS

Foundations that have well-established and published guidelines for their applications and review procedures are relatively easy to access. On the other hand, foundations that have more informal procedures, which involve personal contacts and “networking,” can be a more difficult and mysterious process for many investigators. For this reason, Carlson (1995) recommends that professionals who are seeking funding from a foundation contact members of their own organization’s boards of directors or staff to determine if they know someone in the foundation who can facilitate their initial contacts. Institutional development offices can often serve a similar function.

Working with Development Offices

Large organizations, such as universities and hospitals, often have development offices that are responsible for developing relationships with key staff at selected foundations in order to facilitate the application process and develop long-term funding support for their institution. For this reason, it is clearly in investigators’ best interests to collaborate closely with their institution’s offices of development rather than approach foundations independently. In the interest of facilitating long-term program development, some foundations prefer to work only with staff from development offices concerning applications for support rather than individual investigators. In other instances, an investigator’s institution already may have developed a long-term plan for development of funding from a specific foundation. An individual investigator’s research may or may not fit within this plan. For the

above reasons, a coordinated approach between the investigator and his or her development office is necessary to avoid confusing or frustrating foundation staff and/or alienating the development office at one's institution. Development office staff also may be interested in helping to prepare applications to foundations, which is welcome news for busy investigators.

FOUNDATIONS THAT FUND RESEARCH WITH CHILDREN

Examples of foundations that have targeted psychological/mental health research with children are the William T. Grant and the John D. and Catherine T. MacArthur Foundations. Within the area of research concerning child health and illness, the March of Dimes Birth Defects Foundation, and the Robert Wood Johnson Foundation have funded research focused on chronic conditions. Foundations that focus on specific chronic conditions, such as the Arthritis or Cystic Fibrosis Foundations, also provide potential options for research funding for pediatric and clinical child psychologists, both on a local and national level. While the primary focus of the research that is funded by foundations that focus on specific clinical conditions is generally biomedical, psychologists have been successful in receiving funding from these sources.

Local foundations provide another potential option for research funding that should not be overlooked by investigators. For example, the Cleveland Foundation provided interim support for our research on the outcomes of young children with failure to thrive that was critical to the long-term success of this work (Drotar & Sturm, 1994).

TYPES OF RESEARCH FUNDING THAT ARE AVAILABLE FROM FOUNDATIONS

As is true of research funding that is available from government agencies, a wide range of funding awards are given by foundations. Among foundations that fund research, the number of projects funded and amounts of funding are generally (but not always) less than that provided by government agencies, research funding available from foundations can run anywhere from \$10,000 to \$200,000+ per year. While the amounts of available funding from foundations are smaller than that available from NIH, the applications for foundation research funding are much less detailed than are required for government grants. Consequently, applications to some foundations may not "cost" the investigator as much as government grants in terms of time or energy. In addition to their general grants, some foundations have a category for smaller amounts of funding (e.g., less than \$10,000) that are provided at the discretion of the foundation president and do not have to go through a full board review.

In order to decide whether it is desirable to apply for foundation research support, researchers need to consider whether the level of available foundation funding meets the needs and phase of their particular research program. For some researchers, such as those with relatively well-established research programs, the

level of funding that is offered by foundations may not be sufficient to justify the expenditure of time in filling out the application and preparing reports. For others, the level of support that is provided by foundations may be perfectly adequate to fund pilot or smaller-scale preliminary studies, which are important to develop successful research proposals to government agencies (see Chapter 10, this volume). Alternatively, researchers may be able to utilize foundation research support to fund a specific component to strengthen an ongoing research program, for example, by adding an additional control group that is not supported within existing sources, or to support continued data collection in between cycles of funding from government agencies.

Investigators who apply to foundations also need to consider that the “indirect” or institutional support costs that are provided by foundations to the investigator’s institution are typically much less (e.g., 10%) than are provided by government agencies (often 50% or greater). For this reason, institutions strongly prefer government funding and in some cases their policies prohibit acceptance of the lower indirect cost rates imposed by foundations. However, given the level of competition for government research funding in recent years, most institutions will negotiate their typical indirect cost rates with foundations.

RECEIVING RESEARCH FUNDING FROM FOUNDATION AND OTHER SOURCES

Some investigators may wonder whether it is acceptable to receive from government and foundation sources. The short answer is “yes,” by all means. Because research funding is very difficult to come by, creative investigators will generally leave no stone unturned in their quests for funding. In fact, the research programs of highly successful investigators often resemble “patchwork quilts” of funding that are based on a creative mix of government and private sources. In fact, investigators should understand that foundations are generally much more interested in funding research (or other programs for that matter) that has already received funding from other sources than they are in being the sole source of support for a project. Foundation staff prefer such cost-sharing arrangements because they feel that they are getting greater value for their money. Consequently, it is in an investigator’s interest to have already secured some sources of research support before applying to foundations, if this is at all possible.

The major caveat that researchers should observe in applying for multiple sources of funding from foundations and/or government agencies is that there should not be direct overlap in funding from these sources that support exactly the same activity, for example, the same data collection costs. Such direct overlap of funding is not ethical and is clearly viewed with disfavor by foundations and government agencies. For this reason, researchers who are applying for and/or receiving funding from multiple sources need to carefully consider the potential for overlap in funded activities in preparing their applications and budgets. Investigators should take care to clearly and carefully document the specific activities of their research that are funded by different agencies and/or foundations and identifying how they differ.

FACILITATING SUCCESSFUL APPLICATIONS TO FOUNDATIONS

In order to maximize their chances for a successful application to a foundation, investigators need to ensure that there is a good fit between their research and the foundation's interest and to become very familiar with the foundation's application and review process, which are described in the next section.

Ensuring a Close Fit between the Research Project and the Foundation's Mission

One of the most important considerations in submitting a proposal to a foundation is the degree to which the proposal fits a foundation's stated priorities and mission (see section on finding out about foundations' research priorities). Out of self-interest, some investigators may be tempted to broaden the nature of a foundation's mission and/or priorities to make it fit their research. In my experience, this is not an effective approach because foundations will reject even high-quality proposals that do not fit their priorities. Consequently, prior to their application, investigators need to ensure that there is a reasonable fit between their work and a foundation's interest by careful reading of mission statements and priorities, reviewing the foundation's annual reports, and by making personal contacts and/or written letters with the relevant foundation staff. As is true for applications to government agencies, the time that investigators spend to determine whether there is an adequate fit between their research and a foundation's priorities is well worth their time and energies.

Understanding the Specific Procedures of Foundation Review Committees

Individual foundations use a range of different procedures to review applications. To maximize their opportunities, investigators should become very familiar with the procedures used by the foundation to which they are applying. Some foundations have standing review committees that are similar to the study sections that review research grants for government agencies. For example, the Genentech Foundation for Growth and Development used an eight- to ten-person Board of Directors composed of experienced researchers and practitioners from pediatrics and nursing who serve as reviewers. As with NIH review, each proposal submitted to the Genentech Foundation was assigned a primary reviewer and two secondary reviewers who provide written reviews and a numerical rating (1 = excellent to 5 = not acceptable). Each member of the board participated in the discussion of the proposal and can influence the decision concerning final acceptance or rejection. In some cases, ad hoc reviewers with special expertise were utilized.

Other foundations have less formalized structures and procedures for review and a very different composition of board members and reviewers. For example, many foundations have a very broad representation of professionals (e.g., business and community leaders, philanthropists, etc.) on their review boards or boards of directors and also utilize ad hoc reviewers. The review of proposals for smaller foundations may lie in the hands of several board members, and in some cases, for example, small family foundations, even one or two individuals.

PREPARING APPLICATIONS TO FOUNDATIONS FOR RESEARCH SUPPORT

Specific guidelines in preparing applications to foundations are summarized in Table 1. As with applications to government agencies, researchers who apply to foundations need to read the application materials thoroughly and to follow the instructions carefully. While foundations have highly individual requirements for applications, there are some common themes in the proposed requirements (McIlroy, 1998). Researchers might wish to review the National Network of Grant-makers (1997a) *Common Grant Application*, which contains a sample generic application that is accepted by a number of foundations.

Composing a Letter of Intent

For many foundations, the application process begins with a letter of intent or preproposal, which is a brief summary of the goal and focus of research. Funders may request a letter of intent or preproposal to be mailed to them so that they can screen the proposal content to determine if they have an interest in the proposed research (or clinical service program). It is especially important that this letter be written in a way that clearly communicates the purpose, goals, and significance of the research (Carlson, 1995).

Sample Letter of Intent

To give readers an idea of what might be required in a letter of intent that includes a summary of a proposal, the following is an example of a letter of intent to describe a study of the efficacy of biofeedback for headaches and migraine in children and adolescents. This letter includes the following elements: (1) the need for the program; (2) the applicant's track record and experiences; (3) the goals of the proposal; (4) funding request; (5) plan for outcome evaluation; and (6) anticipated impact and significance. While the following letter of intent format may be somewhat more extensive than is required by some foundations, it will give readers an idea of sample content.

Need

Using headaches and migraine as model conditions, this program will address the need to develop cost-effective management approaches for children and adolescents with chronic pain. Headaches are a very common problem among children and adolescents with estimates of prevalence ranging from 21 to 55%. The prevalence of the most severe juvenile migraine type of headaches ranges from 5 to 10%. Such problems not only occur frequently but have a high level of morbidity as defined by chronic distress, reduced activity, emergency room visits, and persistent school absence. Moreover, many of these children have endured their pain for many years. The factors responsible for the high level of morbidity associated with juvenile migraine and headaches are complex, but relate to such factors as lack of effective treatments within traditional medical care patterns and problems with access to care.

Headaches and migraine in children are often very difficult to manage. Moreover, these problems may not respond to conventional medical treatment, but require comprehensive pain management, including the use of specialized treatment methods such as hypnosis, relaxation therapy, and biofeedback. However, many practitioners are not trained to use these methods with children and adolescents. Even when practitioners have been adequately trained to manage migraine and headaches with appropriate nonpharmacological methods, their management may not be closely integrated with the child's medical care. Consequently, many children with headaches and migraines may not receive the medical help they need.

Based on our clinical experience, for optimal access and impact, nonpharmacological pain management approaches need to be closely integrated with the medical care and management provided by primary care practitioners and neurologists. Also, in this era of managed care, clinical management approaches of headaches and migraines need to be cost-effective and be documented by careful evaluation if they are to have wide applicability. Unfortunately, given the fragmentation of medical care described above, it is not uncommon for children and adolescents who have headaches and/or migraine to consult multiple practitioners for diagnostic workup and treatment, all at considerable expense to families and insurers.

Experiences and Track Record of the Proposed Program

For the past 10 years, staff in the Department of Pediatrics and the Division of Behavioral Pediatrics and Psychology at Rainbow Babies and Children's Hospital in Cleveland, Ohio, have developed comprehensive interventions for children with acute and chronic pain, with a particular focus on headaches and juvenile migraine. This program has several important features: (1) because it is closely integrated with primary care pediatrics and pediatric neurology, it is easier for children to be referred and to receive the care they need at the time that they need it; (2) the program is conducted by practitioners who are experienced in using a wide range of pain management techniques for children; and (3) the program has been effective.

Our clinical experience as well as published studies have indicated that our interventions including biofeedback, relaxation, and hypnosis are effective in reducing the frequency of headaches experienced by children as well as their level of morbidity caused by their pain. For example, in a series of child and adolescent patients ($N = 40$), who were seen over the past year at the onset of treatment, the duration of our patients' headache complaints ranged from 2 to 108 months, with a mean of 25 months. At baseline and subsequent to treatment with hypnotherapy and biofeedback, all patients were rated in terms of the frequency and intensity of headaches. More than 80% of patients reported major improvement as defined by decreased frequency and/or intensity of pain complaints. A fourth (25%) were headache-free, while a minority (less than 20%) reported no improvement. The positive response of children and adolescents to these treatments has led to increased demand for treatment services as well as for education and in-services for primary care practitioners.

Goal of Proposed Program

Funding is requested to build on our program's success by developing and evaluating cost-effective, nonpharmacological approaches to pain management for children with headaches and migraine. One goal of this approach is to treat the majority of headache patients (approximately 80%) within six sessions. We anticipate that our program will provide an efficient approach to the management of headaches, because many children and adolescents use excessive medical resources and medications and often are seen for repeated and sometimes expensive diagnostic workups. Consequently, we believe that utilization of our proposed care path management approach not only will interrupt a vicious cycle of chronic pain for children and families, but also will reduce overall medical costs to families and insurers.

A key feature of the program will be the individualized approach to clinical management. All patients will receive relaxation training, but additional intervention will vary, depending on the nature of the headache complaint. A second key feature of the proposed program will be to extend the numbers of children with headaches who receive nonpharmacological treatments in our community, and thus improve the impact of the program by teaching primary care pediatric practitioners to use methods of relaxation in pain management for selected patients with migraine in their practice.

Funding Request

Funding is requested for a full-time research–clinical assistant and a part-time nurse practitioner who will provide direct services to children and adolescents and who will assist them in training practitioners. Our program will provide in-kind contributions of staff time from the program director and division's research coordinator who will conduct the outcome evaluation.

Plan for Outcome Evaluation

One important feature of our proposed program will involve a detailed approach to evaluation of its effectiveness. The program will be evaluated using a comprehensive plan of assessment using standardized measures that will assess the following outcome:

1. Frequency and intensity of headaches based on self-report measures and parental observation.
2. Assessment of children's functional status and activity level based on child and parent report.
3. Assessment of quality of life as reported by child and parent using the Child Health Questionnaire (Landgraf, Abetz, & Ware, 1996).
4. Patient, satisfaction, and provider satisfaction with care.
5. Description of number of sessions and time spent by providers in intervention.
6. Costs of service provision involving health care costs incurred by families for treatment of their child's headaches.

Anticipated Impact and Significance

The proposed program will develop an efficient and cost-effective model of clinical management of headaches and migraine among children and adolescents that can be generalized to other medical settings. We anticipate that this model of intervention will be of interest to managed care companies because it is a cost-effective method of increasing the quality of life and functional status of children with chronic pain, many of whom undergo multiple medical evaluations and treatments often without success. Our program already has attracted the attention of a local managed care company, which has recently agreed to reimburse treatment using biofeedback for migraines for plan subscribers under the plan's medical benefit. Finally, by training practitioners to utilize relaxation methods in management of juvenile migraine in a wide range of different settings, we will expand the access and availability of treatments for juvenile migraine to those children in our community who need them and practitioners who are trained to employ these methods.

Composing a Formal Application

For some foundations the letter of intent may serve as the initial application. Other foundations have a formal application process and guidelines that specify what should be included in the application. Some of these applications are relatively extensive and rival what may be required for applications to the NIH, while others are less extensive. In the following section, the common elements of an application to a foundation are described.

COMMON ELEMENTS OF RESEARCH PROPOSALS TO FOUNDATIONS

A research proposal to a foundation includes several essential elements. These are: (1) background of the proposal applicant and organization; (2) the significance of and need for the proposed research; (3) statement of the project goals, purpose, and the product; (4) the qualifications of the applicant and institution to carry out the proposed work; (5) method including the circumstances that would facilitate the feasibility of the project; (6) methods of evaluation; and (7) budget information. Readers should consult Chapter 10, this volume, for additional information.

Describing the Backgrounds of the Applicant, the Applicant's Organization, and the Project

Foundations receive many different kinds of grant applications from a great many different organizations. Consequently, one of their first tasks in their review is to understand the applicant and applicant's organization and the nature of the research project. In this regard, reviewers are interested in knowing details about the programs and activities of the applicant's organization, history, and accomplishments and background of the proposed project. For example, foundations may want to know how the proposed project evolved, the current progress, and future directions of the work. Foundation reviewers are very interested in knowing about the investigator and his or her qualifications and experience. Because foundations

do not have standardized formats to give relevant information about investigators, such as the biographical sketch required by the NIH, investigators will need to use their judgment in supplying a sufficient level of information about themselves and their organization that is informative but not overwhelming (see subsequent section on the qualifications of the investigator).

Describing the Need for the Proposed Research

Investigators need to convince the foundation reviewers that the topic of their research is significant, that the research is needed by the field, and may lead to important developments that fit the foundation's interest. The type of documentation concerning the need and significance of a research program that is most convincing to many foundations differs from the primary focus on scientific need and significance that is most convincing to NIH review committees. Foundation review boards are most interested in research that has the potential of furthering the foundation's specific priorities, which may include enhancing general community welfare. For example, in the case of proposed research to fund an evaluation of intervention concerning children's conduct disorders, foundation review committees may want the answers to such questions as: (1) How many children are affected by the problem in question in the community or nationwide; (2) What are the costs to society of children's conduct problems based on indicators such as academic, underachievement, harm to others and property, etc.; (3) What are the costs to society of current treatments, (e.g., juvenile justice system including institutionalization); (4) What are problems with current treatment approaches; (5) How will the proposed research be different than existing research; (6) In the event that the project is successful, how can the information that is generated be used to modify treatment procedures to help children in the community; and (7) Will the project have significance beyond the local community? As is true for any proposal, investigators should not assume that reviewers will regard their project as significant but should carefully establish the significance using concrete examples.

Identify the Goals and Products of the Proposed Research

Foundation review committees will want to identify the specific goals and objectives of the proposed research as well as the concrete products of the research when it is completed. Consequently, the investigator's ability to clearly specify the goals and objectives of the study and what will be learned from the study are critical. Moreover, it is important that applicants to foundations translate their research into concrete products that will be valued by the foundation's review committee. For example, in the case of the hypothetical proposal concerning intervention with children with conduct disorders referred to earlier, a foundation review committee would not only be interested in the scientific product of the research (e.g., the advances in knowledge of treatment efficacy or theory concerning children's conduct disorder), but also in the social impact or significance of the knowledge gained: In the event that the investigator is successful in obtaining new knowledge about treatment outcomes in children with conduct disorders, how could this information be put to use to benefit children and families and reduce the costs to society of conduct disorders? In other words, many foundation review

committees are very interested in the practical benefits to the local community and to society of the scientific knowledge that may be generated by proposed research. Consequently, researchers need to carefully and clearly specify the potential clinical implications of their research to encourage the reviewers to see it as a meaningful project.

Qualifications of the Investigator and/or Applicant Organization to Carry Out the Proposed Research

Foundations are most interested in supporting projects that will be successful. Consequently, investigators should help the reviewers in their task by carefully articulating their own qualifications, special skills, and experiences, as well as those of their co-investigators and research team that will facilitate their ability to carry out the proposed research. If they or their organization has conducted similar research successfully in the past, it will be important to make sure that this experience is clearly stated along with its relevance to the proposed research. Because the foundation board members often are not familiar with investigator's experience, skills, and reputation, the application is the primary vehicle of informing them. Consequently, the information concerning the investigator's qualifications and experience should be complete.

Description of Methods

Investigators who apply to foundations to fund their research should clearly describe the methods of the research, for example, how will their research be conducted (see Chapter 10, this volume). Because the application formats for foundation are generally much briefer (some are between 5–10 single-spaced pages) than they are for government agencies, investigators face difficult challenges in describing their methods to foundation review boards. Moreover, letters of intent for proposals that are part of the initial process of applying to some foundations require even briefer descriptions that are very challenging to write (see previous section).

Given such a very brief format, many investigators will struggle to boil the details of their method down to their essence and clarify their meaning for foundation reviewers without losing scientific rigor. Foundation reviewers are most interested in such questions as how the research will be carried out and what will be done to enhance the feasibility of the project. This very daunting task may require consultation from colleagues and/or in some instances from foundation staff to determine whether the proposed research methods are sufficiently described or whether more or less detail is required. Brief, user-friendly summaries of methods and/or relevant data in tables and figures are especially helpful to reviewers for foundations, who appreciate the clarity of presentation afforded by well-done graphic presentations.

Evaluation of Research Plan

While most foundation review boards do not require the detailed description of methods of evaluation that are the heart of a typical proposal to NIH, they are

nevertheless very interested in how investigators plan to evaluate whether their proposal is successful. Consequently, it is important that investigators describe in nontechnical language the outcomes that will be evaluated in their research, why these outcomes are important, how this evaluation will be accomplished, and identify the benchmarks for success that will be used to answer the critical question: How will the investigator determine that the project is successful?

Preparing Budgets

Foundation reviewers carefully consider proposed budgets. Because these reviewers may or may not have had a great deal of experience in reviewing research-related budgets, investigators need to carefully describe the funds that are necessary to conduct their research and provide justification for these activities. In preparing budgets for foundations, investigators should err on the side of providing too much, rather than too little detail. Some foundations also have specific requirements for preparation of budgets and may request such information as a full year's financial statement and annual operating budget for the applicant's institution, for example, hospital and/or university. If such information is required, applicants should work very closely with their financial officers to obtain the necessary information for the proposal. Finally, foundation reviewers are interested to know the entire costs of the proposed research project, not only the cost of the specific component(s) for which funding is sought. Finally, foundation review boards are also interested in how the investigator plans to fund his or her research when foundation support is no longer in place.

THE FOUNDATION REVIEW PROCESS

As with reviews of research grants by government agencies, the purpose of the foundation review process is to identify those projects that fit squarely with the foundation's mission, that are judged as having scientific merit, yield information that is judged to be significant for the foundation's stated mission, and are feasible and cost-effective. Foundations earmark a certain amount of money to spend for the year to fund a wide range of programs, for example, service, equipment, research, and so forth. Consequently, they seek to identify significant projects that they believe are a good investment. For this reason, it is in an investigator's best interests to make sure that the concrete results and implications of their research are stated clearly and that the scientific product of their proposed research is meaningful to an audience who is not familiar with the methods and potential significance of such research.

Feedback from Foundations during the Review Process

The process of foundation review may differ from what is involved in government agency review in several ways. For example, after the initial review of letter of intent or proposal and prior to full review by the board, a representative of the foundation may contact the investigator for additional information about the proposal and/or order to clarify issues that were not clear. At this point, the foundation

may express interest in some elements of the proposal more than others and may ask for more elaboration of specific aspects of the proposal that are of special interest to them. Consequently, in some instances, an investigator may end up preparing a revised application that is very different from the one initially proposed.

Although a careful response to such requests can be time consuming, it is nonetheless critical because the quality of the investigator's response is carefully judged by foundation staff. Based on my experience, foundation staff are likely to ask the following additional questions: Are there any other research groups in the community or nationally who are conducting similar research? Has the researcher approached any other foundations or agencies about funding his or her research (as noted earlier, many foundations prefer not to be the sole source of funding); If so, what has been the response? What will be the impact of the proposed research on children in the community and their families? How will the research or clinical program be funded after the proposed funds? What are the investigation's long-range goals for the proposed program?

The Site Visit

To facilitate their evaluation of an application, some foundations will require a site visit at the investigator's setting or require the investigators to come to their setting. The purpose of the site visit is for the foundation board or staff to obtain additional information concerning the investigators and the setting in which the research takes place and to answer questions that may have arisen in evaluation of the application. (Similar site visits used to be routinely conducted in reviews of government sponsored research but have been largely curtailed owing to expense.) Some members of foundation boards prefer to have a face-to-face dialogue with investigators in order to gain sufficient information about the proposed project to report back to their fellow board members. To prepare investigators for the site visit, some foundation staff will provide a list of questions and information that they need prior to the site visit. Others prefer to engage in a spontaneous dialogue with investigators and may not provide much information about potential questions. This latter strategy makes it difficult for investigators to anticipate what is to be discussed in the site visit and places demands on them to respond spontaneously and hopefully creatively to questions.

Preparation for the Site Visit

While investigators may not have the opportunity to make extensive preparations for foundation site visits, it is important to anticipate certain events: Because most visits begin with a brief description of the project by the investigator, it is important to prepare a brief summary so that he or she can emphasize the positive "selling points" of the proposal. This presentation also provides an opportunity for investigators to describe the strengths of their proposals. In some instances, a brief slide presentation or handouts for prospective funders can be useful in highlighting critical points of the proposal. Such presentations are most effective if investigators do not simply repeat information that the reviewers already have seen but emphasize unique or additional information in their presentation. In order to prepare for

the site visit, it is important for the investigators to be thoroughly familiar with their own proposal, the major goals and objectives, and the budget. Prospective funders are most impressed with investigators who have a good handle on their own work and who can provide a clear rationale for straightforward questions of fact about their proposal.

In general, site visits are most productive when a spirited dialogue develops between the investigator and site visitors. For this reason, investigators should feel free to communicate their enthusiasm about their proposed projects. Investigators who are shy or terse about describing the strengths of their project run the risk that the site visitors will not convey much enthusiasm to their board concerning their proposals. At the same time, one's enthusiasm for the proposed research and its potential should be tempered with reality. Investigators should remain mindful of the problems and potential weaknesses of their project and have strategies in mind that will facilitate their management of these problems. Finally, it is certainly permissible for the investigators to ask the site visitors about their decision-making process, what will be involved, and the timetable for the review.

Funders are interested in knowing and obtaining relevant information concerning all personnel who are critical to the eventual success of the project. For this reason, investigators should make sure that all such personnel are present at the site visit. Another advantage of the presence of colleagues who will have roles in the proposed project at the site visit is the opportunity it affords investigators to showcase their staff's talents. Other investigators and staff also can provide additional support that can enrich the response to the site visitors' questions. Moreover, the site visitors also gain the opportunity to see how the investigators and staff work together.

Managing Questions in the Site Visit

A general rule for investigators who participate in foundation site visits is to be as factual, open, and enthusiastic as possible when answering questions and engaging in dialogue with site visitors—reviewers. Because they are interested in gathering additional information about the project, site visitors' questions are generally straightforward and are not designed to trick the investigator. In my experience, prospective funders are interested in determining answers to such questions as: What is the major focus of the proposal? What is unique about the proposal? How does the proposal fit with previous work that the investigator has done? How will resources for the project be sustained after the funding that is applied and received is no longer in place?

The principles of responding to site visitors' questions are similar to that of any effective communication. For example, it is essential that investigators make sure that they thoroughly understand the questions that are posed before answering. Even eloquent responses to questions will not impress the site visitors if the wrong question is answered. For this reason, investigators may wish to paraphrase the site visitors' questions whenever they are not clear about them. Moreover, if an investigator does not have the information that is necessary to answer the site visitors question at his or her fingertips, it is permissible to obtain the information and follow-up in writing or by phone.

Feedback Following the Foundation Review

The feedback from foundations concerning research proposal reviews varies from foundation to foundation but is generally much less detailed than reviews from government agencies. In contrast to the detailed pink sheet summary that investigators receive from NIH, the feedback from a foundation review may involve nothing more than a letter informing the investigator that their research project either has been funded or denied. Investigators who receive funding generally are not very concerned about receiving the details of the review. On the other hand, those whose research is not funded are very interested in such information to help them answer such basic questions as: Why was the project not funded? Is it possible or desirable to reapply? If so, what should be changed to strengthen the application? In some cases, investigators may be able to find out more information about their review by contacting the representative of the foundation with whom they are working. Although some foundations will supply more details about the review upon request, others do not, as a matter of strict policy.

TRAINING IN PREPARING RESEARCH PROPOSALS FOR FOUNDATIONS

Most researchers can benefit from specialized training to enhance their capacities to prepare proposals to foundations. While such training can be difficult to obtain, we have had experience with several different methods of training in proposal preparation that may be useful to others, including supervised mentorship and writer's workshops.

Supervised Mentorship in Preparing Proposals for Foundations

In our graduate-level pediatric psychology training program at Case Western Reserve University, supervised mentorship for preparing proposals that are actually submitted to foundations has been a helpful training method (see Chapters 15–19, this volume). Our university is fortunate to have a foundation (the Armington) that funds proposals that focus on research related to children's social-emotional and moral development. Students whose research fits within this category can apply to the foundation for funds up to \$3500 to support their costs for data collection, inter-rater reliability studies, and so forth. Along with those other students and faculty, student proposals are rated by a review group of professors from different departments in the university. Students who apply for funding prepare their proposals under mentorship, which provides opportunity for dialogue concerning strategies of writing, for example, establishing significance, deciding points of emphasis, clarifying complex methods, and so on. Because proposals are limited in length, the process of preparing them are excellent opportunities for students to practice research proposal writing in a highly succinct format. Moreover, a successful proposal provides relatively immediate reinforcement for graduate students in the form of concrete resources that support highly valued activities, such as more efficient completion of their master's or dissertation research.

Under supervision, more advanced graduate students, especially those who are working with pediatric faculty, have developed and submitted collaborative

proposals to foundations to support research. Some of these have been successful. Such collaborative proposals provide not only another opportunity for supervised preparation of proposals to foundations but also a vehicle to discuss research related consultation with physicians (Drotar, 1995).

Recently, I also have utilized individual mentorship to work with junior faculty to help them develop proposals for foundations that support research and/or educational programs. Because application procedures are much less strenuous than for government-sponsored research grants, foundation proposals can be an excellent vehicle for junior faculty to practice their skills in writing research proposals. The process of application to foundations is less cumbersome and generally has a much shorter turnaround time than applications to government grants. Consequently, such applications are generally more feasible than NIH sponsored grants for faculty who are coping with the demands of competing professional activities such as patient care and teaching. In some instances, successful grants to foundations can provide resources to generate data for preliminary studies that can be a stepping stone to government funding.

The Role of Writer's Workshops in Preparing Grants for Foundations

In addition to supervised practice in preparation of proposals to foundations, we also have provided additional training for students in writing proposals to foundations as part of a writer's workshop seminar (see Chapter 16, this volume). In this seminar, students circulate letters of inquiry to foundations for critique by their mentor and student colleagues as a part of the class. Students are then given practice in presenting their research proposals to a mock foundation board composed of the instructor and other students, who ask questions and ask the student to explain or clarify various aspects of the proposal, especially components that were not clear.

Finally, workshops that focus on grant writing to foundations can provide much-needed didactic and experiential training for professionals to learn from experienced grant writers. In my experience these workshops are most effective if they involve a small group and if participants bring proposals in progress that they eventually plan to submit to foundations.

REFERENCES

- Baumgartner, J. E., Feczko, M. L., Hall, V., & Kovacs, R. (Eds.). (1995). *National guide to funding for children, youth and family* (3rd ed.). New York: The Foundation Center.
- Carlson, M. (1995). *Winning grants step by step*. San Francisco: Jossey-Bass.
- Drotar, D. (1995). *Consulting with pediatricians. Psychological perspectives*. New York: Plenum Press.
- Drotar, D., & Sturm, L. (1994). Psychological outcomes of preschool children with early histories of failure to thrive. In B. Stabler & L. Underwood (Eds.), *Growth stature and adaptation, behavioral, social, and cognitive aspects of growth delay* (pp. 221-232). Chapel Hill: Office of Continuing Education, University of North Carolina at Chapel Hill.
- Eckstein, R. M. (1996). *Directory of health grants*. Loxahatchee, FL: Research Grant Guides.
- Fine, T. (1996). Foundations and nonprofit research support. In H. A. Pincus (Ed.), *Research funding and resource manual mental health and addictive disorders* (pp. 127-233). Washington, DC: American Psychiatric Association.
- Freeman, D. F. (1991). *The handbook on private foundations*. New York: The Foundation Center.

- Garonzik, E. (Eds.). (1985). *The foundation grants index* (13th ed.). New York: Foundation Center.
- Grant, A. J., & Bowe, F. G. (1995). Getting the most from federal information sources. In R. M. Eckstein (Ed.), *Directory of grants for organizations servicing people with disabilities* (9th ed., pp. 140–149). Loxahatchee, FL: Research Grant Guide.
- Landgraf, J., Abetz, L., & Ware, J. E. (1996). *Child Health Questionnaire (CHQ): A users manual* (1st ed.). Boston, MA: The Health Institute, New England Medical Center.
- McIlroy, D. P. (1998). *How foundations work*. San Francisco: Jossey-Bass.
- National Network of Grantmakers. (1997a). *Common grant application*. 1717 Kettner Blvd., Suite 110, San Diego, CA.
- National Network of Grantmakers. (1997b). *Grantmakers directory*. 717 Kettner Blvd., Suite 110, San Diego, CA.

10

Preparing Grants to Secure Research Funding from Government Agencies

DENNIS DROTAR

Research funding is critical to the development of research in clinical child and pediatric psychology. The conduct of clinically relevant research in these fields can involve time-consuming and costly tasks involving recruitment and retention of research participants, data collection, and analysis. Moreover, prospective and multisite research projects that are often necessary to advance scientific knowledge in pediatric and clinical child psychology are very expensive and require substantial funding (Drotar, 1994). Research funding is also important to the success of many interdisciplinary research programs, especially large-scale projects (see Chapter 13, this volume).

Beyond the necessity to obtain funding to facilitate data collection for their research, pediatric and clinical child psychologists have other compelling pragmatic reasons to learn how to write grants. For example, researchers in many settings are under increasing pressure to support substantial portions of their salary and research activities through grants. Moreover, a researcher's track record of funding is often critical to his or her career advancement and promotion. University promotion and tenure committees look very favorably on researcher capacities to obtain research funding, because this indicates that their work has been recognized as outstanding in critical review by their scientific peers.

While the need for psychologists who conduct research with children and families to develop skills in grant writing is undeniable, grant writing requires a range of skills that are not easy to master. Moreover, the opportunities to obtain training in these important skills are relatively limited. It is difficult to gain such

DENNIS DROTAR • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

experience and knowledge concerning grant proposal writing in most graduate training programs, which generally do not place much emphasis on the preparation of grants. It is also difficult to acquire grant-writing skills in continuing education courses. Grant writing involves highly specialized skills that require a great deal of practice, ideally under the tutelage of skilled mentors. Consequently, many of us have had to learn whatever craft we have developed as grant writers through on-the-job, trial-and-error experiences in preparing proposals.

My own experiences in writing grants, sometimes securing funding and at other times failing in the process, have generated a number of lessons that are described in this chapter. Readers who are interested in this topic also may wish to consult other sources that provide useful advice to prospective grant-writers (Baron, 1987; Brooks, 1989; Locke, Spirduso, & Silverman, 1987; National Institutes of Health, 1984; Ogden, 1991; Pequegnat & Stover, 1995; Pincus, 1995; Walders, Tanielian, & Pincus, 1999; West, 1997).

PRACTICAL STEPS IN PREPARATION OF RESEARCH GRANT PROPOSALS

Motivation for Preparing a Research Grant Proposal

The preparation of a research grant proposal, especially for a government agency, is a strenuous, time-consuming task. Practical steps in the application process are shown in Table 1. Given the substantial investment of time and energy involved in preparing an application, before deciding to submit, investigators should ask themselves the following questions (hopefully answering affirmatively): "Do I really need to write a grant, and do I have the time, resources, and skills to prepare the application?" (West, 1997). Researchers should carefully consider whether funding from government agencies is absolutely necessary to conduct their research, and if so what type of funding is necessary. For example, while investigators frequently require funding to initiate research, most new or pilot studies will not compete successfully for government funding. On the other hand, a highly appropriate reason for a grant application may stem from the completion of promising pilot work and an interest in expanding the scope of data collection. For example, a researcher who has found positive results for an intervention that has been found to reduce psychological distress among children with cancer at one site might apply for funds to expand the data collection at several sites to obtain greater power and generalizability of findings.

Many powerful motivations may influence the decision to write a grant proposal, such as pressures from one's department or colleagues (Baron, 1987). Many researchers, including pediatric and clinical child psychologists, are under extraordinary pressures to generate increasing portions of their salary through research grants. While such pressures are an essential fact of professional life in many settings, they also may serve to stimulate proposals that are premature, not carefully thought out, and hence destined for failure.

Deciding on Readiness: When Should One Submit a Research Proposal?

One of the more difficult questions faced by investigators is when to submit their research proposals. While a research program needs time to develop to the

point where a successful grant proposal is feasible, economic pressures faced by universities and medical settings give investigators the consistent though often unrealistic message: "the time for grant submission is now." Irrespective of such pressures, investigators need to determine their own readiness based on the following factors: (1) whether they have developed a sufficient track record of previous research that supports their expertise; (2) whether pilot and preliminary studies have been developed to justify the proposal and its feasibility; (3) whether their program of research is sufficiently well developed; and (4) whether the proposal is sufficiently well articulated. Because these questions can be difficult to answer, investigators may wish to obtain consultation from experienced mentors and colleagues, ideally those who have also conducted grant reviews.

Developing Funding for Pilot Work—Preliminary Studies

Review committees appreciate that a mature, well-articulated research program does not spring up overnight. For this reason, a track record of preliminary or pilot studies is critical to establish the basis for the proposed research. Given the critical importance of preliminary or pilot studies in securing research funding, the next logical question is how does one fund such work? Investigators need to be both creative and energetic in identifying resources in their settings that are potentially available to fund pilot studies.

Despite considerable site-specific variation in available resources for pilot and preliminary studies, several options are often possible: First, some universities and departments have earmarked funds to help new faculty develop their research programs. These are usually competitive within the university or individual departments. In addition, funds to support innovative pilot projects may be available from some foundations (see Chapter 9, this volume). Some departments also have resources to support positions for research assistants to run pilot studies and prepare data that might be needed for proposals, for example, chart reviews to establish numbers of patients to project sample size (Drotar, 1995). Finally, investigators in university settings may be able to identify undergraduate or graduate students who are willing to volunteer their time to collect pilot data in exchange for research experience and letters of recommendation.

How to Identify Options for Research Grants with Government Agencies

Both new and experienced investigators struggle with the issue of identifying the most appropriate agency to submit their research. Several strategies may be useful. Investigators might wish to consult the publications of colleagues who work in related areas to see what agencies have funded their work. The *US Government Manual*, which is available at library reference counters or for purchase at the US Government Printing Office (Washington, DC 20402) is a comprehensive resource to identify agencies. Moreover, the *Catalog of Federal Domestic Assistance* published by the Office of Management and Budget researches federal agencies that have assistance programs and grants and allows the user to search for grant programs by subject matter, agency, deadline, date, eligibility criteria, and so on (Grant & Bowe, 1995). Other standard reference works that contain lists of potential sources include the *Annual Register of Grant Support*, *Directory of Research Grants*, *Grant Directory*, and *Foundation Directory* (online at <http://fdncenter.org>).

Table 1. Steps in Applying for Research Grants from Government Agencies

-
1. Determine readiness
 - a. Track record
 - b. Preliminary studies
 - c. Program development
 2. Identify potential options for government research grants
 - a. Reference sources
 - b. Experienced investigators
 3. Contact program staff at the agency to discuss relevance of the application to agency priorities
 - a. Submit summary of proposal to program staff.
 - b. Discuss application and its relevance to the agency or others.
 - c. Decide to submit to a specific agency.
 - d. Obtain consultation from program staff.
 4. Preparing the application—practical considerations
 - a. Read and understand the instructions.
 - b. Begin writing.
 - c. Decide on a realistic timetable for the application.
 - d. Identify key individuals to be involved in the grant writing process/delegate writing responsibilities.
 - e. Decide on consultants for the proposal; convene the group to work on the writing.
 - f. Make assignments.
 - g. Secure secretarial staff.
 - h. Prepare drafts of key sections.
 - i. Follow-up on assignments that are made.
 - j. Leave time for review of the proposal by colleagues.
 - k. Contact program staff to discuss most appropriate study section.
 5. Writing the applications: General guidelines
 - a. Write in clear, concise style
 - 1) Emphasize take-home message.
 - 2) Use headings to highlight information.
 - 3) Use figures and charts to summarize key information.
 - b. Provide a clear rationale for critical methodological and practice decisions.
 - c. Acknowledge limitation of procedures.
 - d. Present alternatives and management strategies.
 - e. Develop a feasible, focused proposal.
 6. Preparing components of the application
 - a. Abstract
 - 1) The central goals of this study and method.
 - 2) The scientific knowledge that will be gained.
 - 3) The long-term implications of the study.
 - 4) The special features of the proposed research.
 - b. Personnel
 - 1) Describe all key personnel that are involved in the project.
 - 2) What are their roles in the project?
 - 3) Carefully justify the time commitments of staff.
 - 4) Obtain letters of support from consultants and contact them.
 - 5) Complete biographical sketches concerning all staff and personnel.
 - c. Budget
 - 1) Make sure that justifications for each item of the budget are well-documented.
 - 2) Make sure that budget requests are closely tied to the project goals and aims.
 - 3) Document any unusual items and/or extraordinary costs.
 - 4) Calculate the time that is required for each of the proposed tasks of the research.
 - d. Resources and environment
 - 1) List all resources and departmental supports that are available to the project.
 - 2) Describe how these resources will facilitate the project aims.
-

Table 1. (Continued)

e.	Specific aims
1)	State aims clearly and concisely.
2)	Provide rationale for specific aims.
3)	State hypotheses.
4)	Describe long-range aims and potential accomplishments.
f.	Background and significance
1)	Critically evaluate relevant research literature.
2)	Describe how the project will address an important need or gap in research on the topic.
3)	Specify what makes the proposed research new and original.
4)	Describe long-term significance of the proposed research.
g.	Preliminary studies
1)	Describe prior research conducted by the research group in the area of the proposed research.
2)	Clearly delineate what has been accomplished in previous research.
3)	Describe how the proposed research has evolved from these preliminary studies.
4)	Clearly describe how the preliminary studies will facilitate the proposed research.
h.	Method
1)	Clarify the basic design of the study.
2)	Describe numbers of subjects, projected sample size, feasibility of obtaining sample size, and power calculation and justification.
3)	Specify all data collection procedures clearly, including what data will be collected where, when, and by whom and at what intervals.
4)	Describe the level of professional background that is needed to collect the data.
5)	Provide a rationale for specific measures, e.g., why are they needed; what is their reliability and validity?
6)	Describe the plan for validation of new measures that are proposed.
7)	Describe the overview of the data analytic plan:
a.)	Justify each step of the data analysis including data reduction.
b.)	Clarify what analysis will be used to complete the project aims.
c.)	State how special data analytic problems will be handled including missing data, etc.
d.)	Justify specialized methods for data analysis.
e.)	Tie the data analytic plan closely to the methods.
i.	Human subjects
1)	Who will be recruited? Why are these particular participants necessary for the proposed research?
2)	Describe the number, gender, and ethnic background.
3)	Carefully describe the risks to subjects and how these will be managed.
4)	Describe assurance that those projects meet the NIH policy for inclusion of research concerning minority subjects and involvement.
5)	Make sure that the project is submitted to institutional review board prior to grant submission.

all of which are available at libraries. Computer search services are available through the National Technical Information Service and the Smithsonian Scientific Information Exchange (Locke et al., 1987). Other useful online resources are Community of Science (<http://fundedresearch.cos.com>) and Grantsnet (www.grantsnet.org). The American Psychiatric Association's *Research Funding and Resource Manual: Mental Health and Addiction Disorders* (Pincus, 1995) contains a detailed profile on federal agencies that fund research related to mental health.

Once having identified potential sources of government funding for their research, investigators should obtain specific information directly from individual agencies concerning the content of their research priorities and directions and a list of currently funded projects. The *Guide for Grants and Contracts* from the

National Institute of Health contains such information. It is available (on-line only) at <http://www.nih.gov/grants/guidelindex>.

In addition to general bulletins that describe their research, many agencies put special calls or requests for application in various program priority areas that they wish to emphasize in accord with their goals for the year. These are listed in the *Federal Register*.

Specific Sources of Government Funding for Research in Pediatric and Clinical Child Psychology

Depending on the content area of their research, pediatric and clinical child psychologists may identify funding opportunities in the following government agencies: National Institutes of Health (NIH) including specific institutes, especially the National Institute for Child Health and Development and various institutes that fund research on specific health conditions, for example, National Cancer Institute, the National Institute of Mental Health (NIMH), especially the Child and Adolescent Disorders Branch, the National Institute of Drug Abuse (NIDA), the Maternal and Child Health Branch (MCHB) of the Bureau of Health Care Delivery and Assistance and the Centers for Disease Control (CDC). Each of these agencies have a complex administrative structure that includes various subdivisions.

To maximize their opportunities for funding, investigators need to stay abreast of continuing changes in the organization of government granting agencies. For example, one such change is the recent restructuring of the NIMH Extramural Research Programs into three divisions: (1) Division of Mental Disorders: Behavioral Research and AIDS, to focus on applied behavioral and social science research including prevention, and early intervention; (2) the Behavioral Science Research Branch, to focus on cognitive, personality, and social processes; and (3) The Division of Services and Intervention Research, which supports large clinical trials of new treatments as well as health services research (American Psychological Association, 1997). Another recent change with important implications is the substantial integration and reorganization of the NIH grant proposal review structure including study sections (Azar, 1998). Investigators should consult the NIH Center for Scientific Review (CSR) web page <http://www.csr.nih.gov> for information.

Types of Research Grants

In addition to identifying the specific agency that is best for their particular research, an investigator needs to consider the type of research funding that fits their needs (Lyman, 1995). Researchers have a large number of options for different categories of research grants within government agencies and should familiarize themselves with these options, some of which are described here. In most cases investigators will submit an application for a research project grant, known as an RO1, which is designed to support a discrete, specified, and circumscribed project that is conducted by the investigator in an area of his or her specific interest and competencies. However, there are other options for research grants, especially within NIMH, for researchers whose work may not yet be competitive for an RO1 include the small grants program, which provides research grant support up to \$50,000 per year direct or project-related costs for up to 2 years for research projects relevant to the mission of NIMH. Priorities for small grants include: (1) research by

less experienced investigators; (2) investigators at institutions without well-developed research traditions and resources; or (3) experienced investigators' exploratory studies that represent significant changes in their research directions or for testing new methods.

Another option for research for new investigators through NIMH is the Behavioral Science Track Award for Rapid Transition (b/start), which provides rapid review and funding decisions for newly independent investigators to submit application for small scale exploratory, that is, pilot research projects related to the mission of NIMH. Support may be requested for up to 1 year with total direct costs not exceeding \$25,000.

A second general type of research award funds investigators rather than specific research projects and again includes several different types of awards. For example, the Mentored Research Scientist Development Award (MRSDA) (KO1) is designed to fund the career development of research scientists who require an additional period of sponsored research experience as a way to gain expertise in an area of research that is new to the candidate or in an area that would enhance the candidate's scientific career development. The Independent Scientist Award (ISA) (KO2) provides support for newly independent scientists who can demonstrate the need for a period of intensive research focus as a means of enhancing their research careers. This has replaced the Research Career Scientist Award (KO5) and provides stability of support (5 years) to outstanding scientists who have demonstrated a sustained, high level of productivity and whose expertise and research accomplishments have been (and will continue to be) critical to the mission of a particular NIH center or institute (NIH Guide, 1990).

A third basic type of proposal is a program project, which is a broad-based, multidisciplinary research program, which has a specific major objective or a basic theme. In contrast to the narrower thrust of the RO1, which focuses on a single project, a program project is directed toward enhancing resources for ongoing research projects, each of which that have a central focus or theme, for example, research on the prevention of human immunodeficiency virus (HIV) infection.

Deciding Whether One's Research Fits with the Goals of a Granting Agency

Because the description of agency priorities is somewhat general, investigators can face a difficult task in deciding whether their research fits agency priorities. For this reason, it is critical that investigators contact agency staff to describe their research program and discuss how it may or may not fit with agency priorities. Such contact will be even more critical in the upcoming reorganization of grant review at the NIH (T. Levitan, personal communication, August 15, 1998). To facilitate such discussions and to allow program staff to render an informed decision about how a proposal might fit, investigators should be as specific as possible about the content of their proposed research. For this reason, submission of a brief summary of the proposal will help program staff to render an effective decision.

What about Applying to Multiple Agencies?

Unless one is working in a highly specialized area, a researcher's work will often fit more than one agency and/or foundation. In order to decide whether to submit a research proposal to more than one funding source, investigators need to

carefully consider an agency's funding limits and priorities, the level of competition for research grants in a particular agency, and track record in funding similar proposals. Because investigators are not permitted to submit the same grant to more than one institute within a particular agency, for example, the NIH, they need to target one agency within the NIH. However, it is possible and in some cases desirable to submit the same or highly similar proposals to different government agencies, for example, the NIH and MCH or CDC and/or research foundations who may be interested in the proposal. Following the principle that the more irons one has in the fire, the more likely it is at least one will stay hot, submitting multiple applications to different agencies is clearly advantageous. Nevertheless, in submitting multiple applications, investigators need to carefully list all applications that are pending and the areas of scientific overlap and differences across proposals.

In the event that similar applications are funded by different sources, investigators will need to inform program staff at these agencies and prepare a new budget that accurately reflects a shared distribution of funding across agencies. Agencies and foundations support such cost-sharing arrangements because they provide opportunities to save money, yet still retain credit for high-quality proposals in their respective portfolios of funded research. Given the level of competition that is involved in obtaining grants from government agencies, investigators should leave no potential funding source unturned and consider alternatives to federal agencies such as foundations (see Chapter 9, this volume).

EVALUATING AND SECURING RESOURCES FOR GRANT PREPARATION

Researchers who are preparing to write a grant need to take inventory of their available resources for this task. In my experience, the critical resources for effective grant preparation are time, didactic training and experience in writing proposals, mentorship and collegial support and administrative support.

Time

Preparation of a research grant to a government agency, especially a successful proposal, is an extraordinarily time-consuming task. In my experience, most investigators, especially inexperienced ones, seriously underestimate the amount of time that is required to prepare a research grant. For this reason, the time that is needed to prepare a grant needs to be carefully considered against the backdrop of the grant writer's multiple professional responsibilities. Ideally, the fact that the grant-writing activity competes with other professional responsibilities should translate into a relatively long lead time between initial preparation and submission, especially if one wants to maximize chances of success. In my experience, many investigators reduce their chances of preparing the best possible proposal by not starting the process of grant writing soon enough. Consequently, they do not leave sufficient time to think through the many decisions that are involved in developing a research design and analytic plan, to obtain consultation from colleagues, to solicit critical reviews of the proposal from others, and carefully to attend to the critical details such as budget preparation. For these reasons, some experienced investigators (at least this one) require what may appear to be a relatively long lead time, for example, 4 months or more, to prepare an effective

application. I have learned from experience that I cannot prepare my best grants under last-minute time pressures.

On the other hand, at times investigators may be pressed into service to prepare a grant application by well-intentioned colleagues or supervisors to respond to the latest request for application that appears to be "tailored made" for their research. However, in my experience, unless the request for application was in fact specifically developed for your project, which is relatively rare but does occur, it is much more effective to develop a proposal to fit ongoing program initiatives.

Efficient grant preparation necessitates having some blocks of time that are reasonably protected from other responsibilities. In order to secure time to prepare a grant application, it may be possible for some investigators to negotiate short-term release time or change in one's responsibilities with one's chair and colleagues. Providing resources for short-term coverage and/or release from activities to allow investigators more time to devote to grant preparation can be a wise long-term investment for program administrators.

Logistical Support

Researchers should identify and utilize any and all potential resources for logistic support for grant preparation that are available to them in their settings. While most universities provide basic information such as the research that agencies fund and the availability of new requests for application, the level and quality of such support varies considerably across universities. However, investigators benefit most from an individualized approach in which they can receive up-to-date information about potential sources of funding that are tailored to their specific research programs. For this reason, researchers need to take the initiative to inform relevant university or agency staff about their research programs and future plans and/or develop their own methods of finding out about opportunities for funding.

Effective grant preparation also requires clerical and administrative support. In my experience, the management of forms and preparation of budgets are very time consuming as well as intimidating, especially for the novice investigator. For this reason, it is very helpful for investigators to have access to both technical support for budget preparation from administrative staff who are experienced in preparing budgets and budget justifications and/or senior investigators, as well as clerical support for word processing.

Researchers who have similar interests also may find it useful to develop ongoing working groups to facilitate preparation of proposals. Such groups, which are especially useful to develop interdisciplinary research proposals, can provide critical support and structure for grant preparation as well as a fertile source of ideas and critique of proposals in process. Because it is very difficult to write a successful proposal by committee, the principal investigator should assume primary responsibility for writing and organizing the protocol to ensure consistency and to clarify roles, responsibilities, and assignments.

UNDERSTANDING THE GRANT REVIEW PROCESS

In order to prepare a competitive proposal, it is important that investigators have a thorough understanding of how grants are reviewed. The grant review process is described in the next section.

Assignment of Proposals to Study Sections

Proposals that are submitted to the Division of Research Grants (DRG) are then assigned to a study section, which is composed of 10–15 reviewers who were chosen by and report to the NIH. Assignments of research proposals to a particular study section are made by examining the proposal content. Because the title and abstract are particularly critical in making this decision, investigators should take some care in preparing these (West, 1997). Although the DRG makes every effort to arrange a suitable match between the content of the proposal and study section, mismatches still occur. For this reason, savvy investigators will try to identify the study section that they feel is the most appropriate to review their proposals. In order to ensure the best possible match between the proposal and the study section, it is in the investigators' best interests to be in contact with program staff to discuss potential review committees. Investigators can request a particular study section to review their proposal in a cover letter that is submitted with their application. Program staff also can help investigators prepare their abstracts to ensure that the grant proposal is assigned to the most appropriate study section.

Investigators should recognize that even with careful and proactive preparation, one does not always obtain a precise match between study section and proposal content, owing to the combination of tremendous scientific specialization coupled with breadth of the field, which is reflected in the extraordinary range of proposal content. Baron (1987) notes that a lack of fit in the review process is more likely to occur if a proposal falls on or near the boundary of two separate fields, in a new and innovative area, or is an area that is investigated by a small number of scientists. In order to address specialized areas of content, ad hoc reviewers may be appointed to the study section by the executive secretary based on the investigators' suggestions. Researchers who feel that adequate review of their proposals may require special expertise from ad hoc reviewers should not hesitate to make such a request.

Communication with Agency Staff after the Proposal is Submitted

Within 6–8 weeks of their submission, investigators will receive notification of the review committee or study section to which their grant has been assigned the study section's contact number. The executive secretary is the contact person with whom investigators should communicate after their grant is submitted. Investigators are strictly prohibited from discussing their application with members of the review group. The executive secretary has the difficult job of setting the agenda, assigning reviewers, fielding inquiries from investigators, and distributing supplementary material for review that can be submitted after the deadline. Such supplementary information should be brief and highly relevant to the grant review, for example, critical errors that were noted after the grant was submitted, brief summaries of new data analyses and/or updates, or manuscripts based on the proposed research that were submitted, accepted, and/or published after the grant was submitted.

Study Section Structure and Operation

Study sections are composed of senior scientists with established reputations in their respective fields, and publications and funded grants in content areas that

are relevant to the priorities of an agency. Members of study sections are appointed for 4 years and are chosen to represent various areas of scientific knowledge, for example, health promotion and/or methodological, for example, statistical models of longitudinal data analyses that represent the interests of an institute or agency. They are nominated by the executive secretary of that section based on scientific expertise and willingness to serve. Policy considerations such as minority representations, geographic distribution, and professional discipline are also taken into account in putting a study section together. Nevertheless, investigators who conduct research with pediatric and clinical child populations need to appreciate that special expertise in these specific areas are infrequently represented on study sections.

Given the importance of the review process, investigators who accept an appointment to a study section must be dedicated to this enterprise. Responsibilities include three meetings a year for 2–3 days, spending time reviewing grants prior to the meeting, developing written critiques, and presenting this information at the meetings at a nominal remuneration (\$150.00 per day), in addition to travel expenses. Potential motivations to participate in the difficult chores of a study section involve prestige, the opportunity to gain firsthand experience with the review process, collegiality, and the opportunity to perform a service to the field.

Study section members are assigned to review (as primary, secondary, or tertiary reviewer) grants that the executive secretary judges to be in their area of expertise. Reviewers are also assigned grants based on their general expertise, for example, in statistical methods, rather than their knowledge of a scientific content area. Depending on the number of grants that are received for a particular cycle, each member of the study section may review three to four grants as primary and a similar number as a secondary and/or tertiary reviewer (West, 1997).

Review Criteria

The reviewers are given explicit instructions concerning what to look for in evaluating the grant, preparing their written critique, and assigning their score. The following is a text of the most recent revision in review criteria which are applicable to all grants submitted on or after October 1, 1997 (Agnew, 1997; Grant, 1997). Reviewers consider five criteria in assigning the overall score for a proposal but an application does not need to be exceptional in all categories to be awarded a high priority score. The criteria are as follows:

1. *Significance*: Does this study address an important problem? If the aims of the application are achieved, how will scientific knowledge be advanced? What will be the effect of these studies on the concepts and methods that drive the field? How will scientific knowledge be advanced?
2. *Approach*: Are the conceptual framework, design, methods, and analyses adequately developed, well-integrated, and appropriate to the aims of the project? Does the applicant acknowledge potential problem areas and alternative tactics?
3. *Innovation*: Does the project employ concepts, approaches or method? Are the aims original and innovative? Does the project challenge existing paradigms or develop new methodologies or technologies?

4. *Investigator:* Is the investigator appropriately trained and well suited to carry out this work? Is the work proposal appropriate to the experience level of the principal investigator and other researchers, if any.
5. *Environment:* Does the scientific environment in which the work will be done contribute to the probability of success? Do the proposed experiments take advantage of unique features of the scientific environment or employ useful collaborative arrangements? Is there evidence of institutional support?

In addition to the above criteria, in accord with NIH policy, all applications also will be reviewed with respect to the following areas: (1) the adequacy of plans for recruitment and retention of subjects to include a broad representation of genders, ethnic minority groups, and children as is appropriate to the scientific goals of the proposed research; (2) the reasonableness of the proposed budget and duration in relation to the proposed research; and (3) the adequacy of the proposed protection for humans, animals, or the environment, to the extent they may be adversely affected by the proposed research.

Review Process At Study Section Meetings

Investigators need to understand the process of grant review at study sections. The review typically begins with the primary reviewer's oral presentation of his or her written critique of the grant followed by the secondary reviewer and tertiary reviewer. In addition to noting the proposal's strengths and weaknesses, each reviewer also makes a recommendation for acceptance or rejection. Other committee members then add their comments and questions in a brief discussion. Study section members are strongly encouraged to recognize potential conflicts of interests and excuse themselves from reviewing the work of colleagues from the same institution, or those with whom they have a collaborative relationship or a significant conflict.

Given the volume of grants that need to be reviewed, study sections cannot spend a long time on any individual grant. Because study section members cannot review every grant in detail they can rely heavily on the assigned reviewer's opinions. For this reason, grants that stimulate a strong consensus (either in favor of acceptance or rejection) are dispatched quickly, while those that provoke divergent opinions generate more discussion. In my experience, disagreement among study section members concerning the merits of proposals are comparable to that which occurs in manuscript review, for example, different reviewers give different weights to the components of the proposal, have different standards for aspects of method and design, or may notice different flaws or strengths.

Assignment of Scores

Following the presentation and discussion of a proposal, study section members then cast their ratings via secret ballot. If a proposal is recommended for funding, reviewers assign a score, ranging from 1 to 5 with a rating of 1–1.5 being outstanding; 1.5–2, excellent; .2–2.5, good; 2.5–3 satisfactory; 3–3.5 adequate; 3.5–4, poor; and so on. Investigators should understand that it does not take very many methodological problems for reviewers to lower their rating down from a 1,

which is highest, to a 2 which is still very good, but clearly lowers the probability of eventual funding. Moreover, a proposal's rating may be raised by the enthusiasm of an individual reviewer who decides to "champion" the grant or lowered by one who is very troubled by a particular methodological problem. Because each study section member is responsible for presenting a critical evaluation of proposals to their colleagues, there is pressure to detect significant methodological or feasibility problems, lest such flaws be uncovered by their peers. As seasoned grant writers will attest, this feature of the review process makes it very difficult for all but the most well-written, articulated, and well-designed grants to escape this gauntlet of criticism relatively unscathed.

Following the review, a grant receives a numerical score that is composed of averages of the ratings of scientific merit multiplied by 100 from all the study section members. Such rankings are used to balance differences in ratings among and within study sections. Investigators also receive a brief summary of the reviewer's conclusions written by the executive secretary and critiques that were written by the primary, secondary, and tertiary reviewers. This brief summary includes the following elements: (1) description of the proposal; (2) reviewers' critiques; (3) gender, minority, and human subjects considerations; (4) budget issues; and (5) priority score and percentile rankings. Details of the review are not discussed with investigators until after the summary has been sent out. After they receive the review summary, investigators can call the executive secretary to clarify aspects of the review that were not clear and for advice concerning revision and resubmission.

Funding Decision: The Role of the Advisory Council

Following review study section review, applications are then considered by the advisory council of the particular institute to which it is directed. Percentile rankings for all proposals are rank ordered to form a single list. Agencies generally allocate funds on the basis of these scores. The lower the percentile priority score, the greater the likelihood of funding. Depending on agency and funding, the ranges of fundable scores for research projects related to pediatric and child clinical psychology may range from 100 to 150 or higher. The specific cutoff score that is needed to achieve funding will depend on the amount of funding allocation that the agency receives for the particular year and the level of competition, as defined by the number of grants that have been judged by the review committee with potentially fundable scores.

The award of funding is not made until the agency has received notification of its funding allocation for the current year and the advisory council of this agency has approved the funding decisions for that particular grant cycle. The purpose of the advisory council review is to consider funding specific grant applications in light of institute priorities and reduce duplication of funding (Matthews, 1997). Generally, the advisory council will concur with the priority score recommendations that were made by the study section. Nevertheless, the council will occasionally recommend that certain proposals be given special consideration based on their relevance to the institute's mission and if they concern topics that are underrepresented in the institute's portfolio. The final step before awarding funding is to determine the actual size of the award. At the NIH, budgets are generally cut

(typically 10% or more) in the eventual award of funding. In some cases, very expensive applications that were rated highly by the study section can be viewed by the council as too expensive relative to their benefits for public health (Matthews, 1997).

Resubmissions

In my experience, grant proposals are rarely accepted by government agencies on the first review. One reason for this is that it is difficult if not impossible for investigators to anticipate and address the specific concerns of a study section. After they receive the funding decision, investigators will want to carefully review their summary statements, sometimes called pink sheets, because they were formerly written on pink paper (Lyman, 1995). An applicant whose grant is not funded has several basic options: (1) revise and resubmit the application; (2) revise and apply elsewhere; (3) accept defeat and prepare an entirely different application; and (4) appeal the judgment. In deciding whether to resubmit a proposal, an investigator needs to determine whether the methodological and feasibility problems that were noted by the reviewers are correctable with a revised application. This can be a very difficult decision that usually requires consultation from colleagues and/or the executive secretary of the review committee.

Investigators are rarely satisfied with a review that results in the rejection of his or her grant and may question certain elements of the review, for example, comprehensiveness or the weight that the committee gave to particular aspects of their application. In exceptional circumstances, that is, if investigators feel that the review committee has seriously misconstrued critical elements of the proposal and/or has rendered an unfair or biased review, they have the right to appeal the review by submitting a detailed letter rebutting the review to the program administrator of the relevant institute. If the council feels that the application's objectives have merit, it may recommend that the application be deferred and re-reviewed. The conduct of an appeals case takes at least 4 months.

Recent Modifications in the Grant Review Process

In an effort to streamline and improve the efficacy of the review process and to increase the amount of time that the review committee can devote to potentially competitive proposals, the NIH recently has introduced several modifications in the review process. The one most significant modification is the introduction of a triage system by which grants are given an expedited review by one or more members of the committee prior to the committee meeting. The grants that are judged to be noncompetitive are rejected and not reviewed by the entire committee. As is the case for traditional review procedures, applicants still receive a pink sheet summary and have the option to reapply.

HELPFUL HINTS FOR PREPARING A GRANT APPLICATION

This next section summarizes information that I have found helpful in preparing successful research proposals to government agencies such as the NIH, NIMH, and MCH. Readers also may wish to consult several other sources that contain other

tips for preparing grant applications (Baron, 1987; Bonawski, 1996; Brooks, 1989; National Institutes of Health, 1984; Pequegnat & Stover, 1995; West, 1997). See Table 1.

Read and Understand the Application Instructions

In order to avoid preventable mistakes in their applications (Baron, 1987), the investigator's first task is to carefully read the entire set of instructions and application forms for the grant application. This seemingly straightforward task may take multiple rereads, especially if it is the investigator's first application. Because it is possible to misconstrue these instructions, investigators should use colleagues and informed grant specialists to interpret the instructions that are not clear.

Develop a Realistic Timetable to Prepare the Proposal

In order to plan one's time effectively and break down the tasks of grant preparation into manageable components, prospective grant writers should prepare a timetable for when various parts of the grant application, for example, drafts of various sections, budget, and so on, should begin, target dates for completion, and assignments of responsibilities for preparation.

Investigators should familiarize themselves with deadlines for grant submissions, which for the NIH are three times a year for new proposals: February 1, June 1, and October 1. Researchers, most especially those with ongoing research projects, need to be mindful of the timetable for the review process, especially the lag time between submission and funding. For example, grants that are submitted for a February 1 deadline are not reviewed until June or July, with the earliest start date of December 1. Given the lag time between submission and funding, investigators need to plan the timeline for their proposals carefully, especially for applications to continue ongoing funded research, which are known as *competing continuation grants*. Investigators also should familiarize themselves with their university/hospital procedures for processing grants, for example, steps for processing of budgets, securing relevant signatures, and so forth. Consequently, components of the application, especially the budget, will need to be ready for processing in advance of the grant deadline.

Begin Writing

Initiating the commitment to the proposal by beginning writing in earnest is a critical step in grant preparation. In order to build some momentum in their grant writing, investigators may want to begin with a section of the proposal that is the most comfortable for them. Starting with the literature review is a comfortable choice for some researchers because this is a familiar task. However, grant writers should be wary of spending too much time on the literature review, which is only one component of a proposal, and not leaving enough time for the most critical piece, which is the proposal itself.

Delegate Tasks

In developing one's plan for proposal writing, it is useful to identify the key individuals who will be involved in preparing various components of the proposal. Wherever feasible, investigators should delegate some of the processing of compo-

nents of their proposal to colleagues and/or support staff. In fact, certain aspects of a proposal, such as preparation of investigators' biographical sketches, need to be delegated to others, and other research support.

Obtain Consultation Whenever Possible

Over and beyond their role in informing investigators about their funding priorities, program staff at government agencies may be available to provide consultation to investigators by reviewing drafts of their applications in order to address such questions as: Does the application effectively address the priorities of the agency? Are there glaring omissions of content? What information needs to be added or clarified? Investigators should ask program staff about their availability to provide such consultation, which can be very helpful though not sufficient, in my experience (Lorion, 1995).

To maximize their chances of funding, investigators also should obtain consultation from colleagues who have had substantial (and hopefully, successful) experience in preparing grant applications and colleagues whose expertise is relevant to specific areas of the application, for example, specialized statistical analyses, measurement, issues, and so on. Inexperienced grant writers clearly benefit from information concerning the "tricks of the trade" from more experienced investigators who can identify ambiguous writing, as well as relevant details and/or critical methodological issues that may have been overlooked. In order to make best use of such consultation, would-be applicants need to leave sufficient time to allow for outside review and revisions based on this advice. For this reason, investigators who are fond of the "last-minute" style of writing place themselves at a considerable disadvantage relative to their more organized and planful peers.

Write in a Clear, Concise Style

The strenuous demands on reviewers to process large quantities of highly technical information in a short period of time heighten the necessity for clear and concise grant writing. (See previous section on overview of the grant review process.) Even mild-mannered reviewers can become frustrated (if not downright annoyed) by the labor involved in trying to decipher an obscurely written proposal and identify key information that should be easily accessible in the application, but is not. Clarity of writing also is very important because reviewers may not be familiar with the specific content area of the proposal.

Investigators can enhance the clarity of their proposals by structuring their proposals around specific take-home messages that are highlighted in the text by headings and subheadings and by using figures, charts, and tables in key areas of the proposal that may be difficult for reviewers to understand. Sampling procedures, and so forth, include theoretical frameworks, causal models, expectations for hypotheses, sample selection/recruitment procedures, and time limits and the anticipated flow of data collection (see Pequegnat & Stover, 1995, Appendix C).

Provide a Rationale for Key Methodological and Practical Decisions

Reviewers are interested in evaluating the quality of the investigator's problem-solving ability to make informed decisions concerning key scientific issues, for

example, the data analytic plan and practical problems such as determining the personnel that are needed to implement the project. Reviewers should be able to follow the investigator's logic in dealing with the key problems of their research proposals. Consequently, investigators should specify the rationale for their decisions so that reviewers can appreciate why specific research designs, procedures, or methods of analysis were chosen over available alternatives.

Acknowledge the Limitations of Procedures and Measures

Proposal writers should acknowledge potential limitations in their methods and the possible impact on their findings and state how these problems will be managed. Reviewers appreciate the fact that every study, especially research with clinical populations, has inherent limitations. Nevertheless, they want to know whether the investigator has carefully considered problems that arise in measurement, data analysis, or implementation and how an investigator plans to manage them. Investigators' ability to anticipate critical methodological problems, effectively defend the proposed methods and say why they were chosen over alternative methods, and develop effective strategies to address problems counts heavily in their favor. Alternatively, if a reviewer uncovers significant problems that an investigator has not considered, this is taken as a sign of naiveté and will be weighed heavily in the critique. While relevant methodological limitations need to be addressed, investigators need to be selective and focus on the most important problems, lest their proposal be viewed as seriously flawed.

Integrate the Components of the Proposal

To be most convincing, investigators need to assure reviewers that the various components of their grant proposal fit together and are internally consistent. Because grant proposals involve an extraordinary amount of material that is generally written in stages, such integration can be difficult to achieve. For example, one common problem is to mention a large number of measures and variables in the method section but fail to describe in detail how each of them will be analyzed. To ensure that their proposals are well integrated, investigators should develop central themes that are reiterated throughout the proposal. Consultation from colleagues who are instructed to review the proposal to determine how well the different sections are integrated can enhance the quality of proposals.

Develop a Feasible, Focused Proposal

In their zeal to make their proposal as attractive to reviewers as is possible, some researchers propose projects that have a very broad scope (e.g., multiple data points, multisites) and address multiple questions. Unfortunately, such complex, multifaceted projects not only are difficult to accomplish but they are difficult to describe effectively in the constraints of a grant proposal. The larger and complex the proposed project, the more difficult it is for investigators to describe methods and analytic procedures clearly and completely for each facet of their proposals. For this reason, investigators should limit the scope and focus of their proposals so that not only is their research perceived by reviewers as clear, but feasible.

Consequently, it is preferable for investigators to tackle fewer questions (Cohen, 1992) and elaborate their methods in great detail (Cohen, 1992) than to propose a large number of questions and deal with them superficially.

SUGGESTIONS FOR PREPARING SPECIFIC COMPONENTS OF THE APPLICATION

Overview of the Application

The first section of a grant application contains the face page, which identifies the title of the project, institution, investigator, and costs; page 2 includes the abstract and listing of key personnel; page 3 lists the budget for the first 12 months of the grant. Other components of the first section of the proposal include the proposed budget for the entire project period, biographical sketches for project investigators and other personnel, a section on other grant support received by the investigators, and a description of resources and environment.

The second section of a grant application contains the research plan, which includes the following sections: (1) specific aims; (2) background and significance; (3) progress report—preliminary studies (a progress report is required for competing continuation and supplemental applications; preliminary studies are described as useful but optional for new applications); (4) experimental design and methods; (5) Human Subjects; (6) vertebrate animals (if applicable); (7) consultants—collaborators; (8) Consortium—contractual arrangements (if applicable); and (9) literature cited. Most investigators also will submit an appendix, which includes such information as relevant publications, manuscripts accepted for publication, and other printed materials (e.g., tables, charts, technical reports illustrating methods, etc.) that have resulted from the proposed project.

Abstract

Because the abstract is the very first part of the proposal that the reviewers see, it can set the tone for the review. Consequently, abstracts should be clearly written and describe essential study goals (Locke et al., 1987), carefully describe the scientific knowledge and/or long-term implications of the proposed project and new or special features of the proposal. In developing their abstracts, investigators might wish to select key, carefully crafted phrases from the specific aims, background, and/or method of their proposals. Consequently, it may be most effective best to polish the abstract after the proposal is well developed or completed.

Personnel

An experienced, well-chosen research team is critical to the success of a proposal. The key personnel for the proposed project are listed on the abstract page, described in the proposal, and in individual biographical sketches. As experienced scientists, members of study sections recognize that research projects require a wide range of professional expertise. For example, depending on their specific focus, research proposals with pediatric and clinical child populations might

involve investigators with expertise in pediatrics, developmental psychology, child psychiatry, and/or statistics, and so forth.

However, investigators need to carefully evaluate whether and how the quality of their proposals would be improved by the addition of co-investigators or consultants. Some investigators, especially novice researchers, make the mistake of proposing highly sophisticated statistical methods, such as, structural equation modeling, or specialized assessment methods, such as, security of attachment, but do not include personnel who are experienced with such procedures in their proposal. This mistake is easily detected by reviewers. On the other hand, it is equally problematic to include a large number of personnel, especially coinvestigators, in a proposal without adequately justifying their unique roles in the proposed research. Consequently, the specific contributions and expertise of specific personnel always need to be carefully weighed and justified against their costs.

Another common problem in proposals is the failure to sufficiently justify the personnel that are listed, to document their expertise, and justify their specific time commitment. For this reason, it is important that investigators discuss their proposals with potential coinvestigators, identify their specific contributions and roles, and carefully plan their time commitments prior to submitting the proposal.

Ideally, personnel who are described in a proposal should have a working relationship and history of collaboration, because this is most convincing to reviewers. Wherever possible, it is also important to identify specific individuals for each position and include their biographical sketches. For this reason, investigators who have access to a wide range of possible coinvestigators and consultants have a clear advantage over investigators who work in isolation.

Investigators who wish to include consultants to help them implement specialized analytic methods or measurement tools (e.g., using structured interviews to obtain child psychiatric diagnoses) should contact them prior to the proposal submission to discuss the project and their availability and commitment. Reviewers like to see evidence that the investigator has discussed the project with the proposed consultants and that they have agreed to participate. Consequently, letters of support should be obtained from consultants that document their interest and commitment to their role in the project and included with the proposal.

Biographical Sketches

Reviewers rely heavily on up-to-date biographical sketches of project investigators and staff to document their qualifications and capabilities to carry out the proposal. The directions for preparing these sketches state the essential content areas, including previous employment, experience, honors, professional activities, references for publications in the past few years, and representative early publications. Junior investigators should make sure that all their training and professional experiences, especially those that are relevant to the grant, are listed in their biographical sketch. However, clearly it is not advisable to pad biographical sketches with irrelevant professional experience or manuscripts in progress. On the other hand, experienced investigators have the “problem” of selecting and highlighting those particular experiences and publications that are particular relevant to the content of the proposal. For this reason, some investigators may wish to

design several biographical sketches to allow them to emphasize specific research experiences that are tailored to different proposals.

Preparing the Budget

One of the primary tasks that reviewers are charged with is to determine whether the proposed resources are adequate to carry out the proposed project. If they are not, the proposal may be judged as not feasible.

Many reviewers also regard the quality of budget preparation as an important indicator of an investigator's skills in managing a research project. In fact, some reviewers will look at the proposed budget first in order to determine the investigator's priorities and to evaluate how the investigator has planned to allocate resources that are necessary to implement the project. See Mucha (1995) for a description of dos and do nots in preparing budgets, and Pequegnat and Stover (1995, Appendices B & C) for a sample budget description.

Reviewers are asked to review proposed budgets carefully to determine whether the resources that are proposed are sufficiently justified, essential to the project aims, and in line with customary costs. Because grant reviewers are experienced investigators, they generally are very knowledgeable about research-related costs, and hence are not terribly forgiving when they uncover projected costs that are clearly out of line with reasonable expectations (Baron, 1987).

In preparing their budgets, investigators need to be as realistic as possible in estimating the costs of running the proposed project and to carefully document how they arrived at these estimates (Baron, 1987). This seemingly straightforward guideline is difficult to implement, especially if one has not actually previously conducted a study that is similar to the one that is proposed. For this reason, in preparing their budgets, investigators should seek advice from senior investigators and/or knowledgeable grant administration staff and look at budgets from successful proposals.

Personnel needs consume the majority of research costs and should be carefully anticipated. To estimate their personnel needs, investigators should estimate the time that is necessary for data collection, preparation and analysis, coordination of the project, and so forth, and describe how the various responsibilities and time commitments will be allocated across different staff. Unfortunately, such time requirements are difficult to estimate, partly because they change in response to different phases of the study (see Chapter 11, this volume). Investigators and coinvestigators should budget sufficient time to allow them to conduct the research. Reviewers are quick to question whether an investigator's proposed commitment is sufficient to allow him or her to conduct the research.

For each aspect of their proposed budgets, investigators should provide very detailed justification of why the projected expenditures are necessary and how they were derived. These calculations should be sufficiently detailed to allow reviewers to follow how the projected costs were derived. For example, to estimate time for research assistants who conduct outcome assessments, investigators should document how much time will be needed to arrange and conduct each assessment and multiply by the number of assessments and participants in the study. One final note of caution: Budget padding is not a good idea because it is often detected and calls the investigator's judgment, if not integrity into question (Baron, 1987).

Resources and Environment

Reviewers want to know whether an investigator has access to physical resources (e.g., laboratory and office space, computer support) that will be necessary to carry out his or her proposed research. Consequently, researchers should be careful to list all resources that are realistically accessible to their proposed projects. In the course of taking such inventories, researchers may discover that they do not have sufficient resources (e.g., office and laboratory space, necessary equipment) to conduct their projects in the event that they are funded. In such cases, prior to their grant submission, investigators should ideally negotiate with relevant administration concerning the resources to conduct their project (see Chapter 11, this volume). However, in practice, space commitments are not made until the investigator has actually secured research funding.

Research Plan: Overview

The research plan, which is the heart of the proposal, requires investigators to address four basic questions in detail: (1) What do you intend to do in your research; (2) why is the work important; (3) what has already been done in the area of the proposed research; and (4) how is the work going to be done? Readers should consult Pequegnat and Stover (1995, Appendix A), to see a research plan for a successful proposal. The elements of the research plan are described in the next section.

Specific Aims

In the specific aims section, investigators are asked to describe concisely and realistically what their proposed research is intended to accomplish, the specific hypotheses that will be tested, and long-range objectives, all within the one-page recommended limit. One common mistake in the specific aims section is to state very general hypotheses or hypotheses with an unclear rationale (Dawes, 1995). Highly specific descriptions of study aims are prized by reviewers, while abstract, rambling discourse is not. Moreover, aims should be clearly tied to hypotheses.

Background and Significance

In the two- to three-page limit for the background and significance section, investigators need to accomplish the following: (1) describe the background to the proposal; (2) identify the specific gaps in the scientific literature which the project is intended to fill; and (3) concisely state the importance of the research by relating specific aims to longer-term objectives. The background and significance section is a delicate blend of a well-reasoned, scholarly critique of the literature and promotion or advocacy for the proposed project in which the investigator highlights the significance of the proposed project to reviewers.

The essential challenge of the background and significance section is to make a clear case for the importance of the work to reviewers who may not be familiar with the proposed research and/or may not share the investigators' enthusiasm for the importance of their studies. Investigators need to accomplish three basic tasks in this section: (1) critically evaluate previous research (their own as well as others);

(2) identify the specific gaps in scientific understanding that their research will address; and (3) state why obtaining this information is important. A project's funding can depend on whether the investigator can convince the reviewers that (1) the proposed research has not been done (or has not been done well previously); and (2) needs to be done to advance science.

Establishing the Significance of Research

To carefully establish the significance of a research proposal, investigators should address several important questions: What are the specific features of the work that make it new and interesting (e.g., a novel theoretical approach, approach to measurement, research design, etc.)? What is the anticipated product of the proposed research? What is the scientific significance of the proposed work (e.g., how will it advance the field?).

Consider an investigator who proposes a study of the effects of a new family-centered approach to improve treatment adherence among adolescents with diabetes. He or she must first establish that there is a scientific need for such a study in this population. This can be accomplished by documenting that problems with adherence occur frequently in adolescents with diabetes and that noncompliance may affect their short- and long-term health outcomes. Moreover, the investigator also should make it clear that few studies have carefully assessed the impact of psychological interventions on adherence to treatment. One could also argue that prior studies of intervention have been limited because they have not addressed potentially important theoretical issues, for example, the relationship between family communication and adherence, that will be the focus of the proposed research [see Herek (1995) concerning theoretical frameworks for research proposals]. Finally, this investigator could then articulate how the results of the proposed intervention and the principles that underlie the intervention model (e.g., sensitivity to critical developmental and family issues) might apply to adolescents with diabetes as well as to other populations of adolescents with chronic health conditions.

Reviewing Research

An effective background and significance section should summarize and highlight previous research that is relevant to the proposal, including work that has been done by the investigator. Given space constraints, such a review should be highly selective, focusing on the specific questions that will be addressed by the proposal and articulating the specific scientific issues that are directly related to proposed aims.

Reviewers are most impressed by a scholarly critique of previous research that carefully articulates the scientific import of the proposed study by addressing the following questions: What new approach, model, measures, and so forth, will the investigator use in the proposed study? In what ways will the proposed method extend previous work? What methodological and/or theoretical problems will be addressed?

Some novice investigators make the mistake of describing such a large number of methodological problems that reviewers may come to believe that the proposed

research is inherently flawed. Investigators should not imply that their studies will solve every methodological problem that is described, lest their proposed research be held up to an unintended and impossible standard.

Preliminary Studies—Progress Reports

In the preliminary studies section, investigators should describe research that will help to establish their experience and competence to carry out the proposed study. Although the grant instructions lists preliminary studies as an option for new research proposals, they are in fact essential. Reviewers weigh investigators' previous research track record very heavily for several reasons: Previous studies help to articulate and refine research questions and hypotheses, provide evidence that the investigator is knowledgeable about the proposed research, can help to establish the reliability and validity of measures, and document the feasibility of the proposed project.

Reviewers are understandably partial to research that is a logical outgrowth of an ongoing research program compared with proposals that lack evidence of continuity and feasibility. The best way that an investigator can gain credibility and convince reviewers of the feasibility of their proposed research is to demonstrate that he or she has done and/or is currently doing such work. This is important for any investigator, but is especially so for junior investigators, who ordinarily do not have an extensive track record of previous research.

Detailed progress reports are especially important for renewals of research funding. Reviewers need to be convinced that the currently funded research project has led to significant findings, that the investigator has been productive in publishing them, and that the project is likely to lead to additional scientific discovery and publications. In summarizing their research progress, investigators should highlight the scientific significance of new findings rather than leaving it to the reviewers to discover such information. In considering applications for renewals, reviewers critically evaluate whether the investigators have achieved their stated aims and have accomplished what they set out to do (and more) and/or have provided a detailed justification for changes in methods or less-than-optimal progress. Finally, the data that are described in a progress report ideally should lead directly to and hence set the stage for the proposed studies.

Experimental Design and Method

In no more than about 20 single-spaced pages, investigators are instructed to describe the experimental design and procedures that will be used to accomplish the specific aims of the project including the means by which data will be collected, analyzed, and interpreted. Moreover, they are asked to describe any new methodology and its advantage over existing methodologies, the potential difficulties and limitations of the proposed procedures, alternative approaches to achieve study aims, and a sequence or timetable of tasks. Finally, investigators are asked to point out procedures, situations, or materials that may be hazardous to personnel and the precautions that will be exercised.

Clarity of Basic Study Design and Data Collection

Reviewers need to understand the basic design of the proposed research so that they can evaluate how well it is suited to accomplish the primary research aims. Many projects in pediatric and clinical child psychology involve multiple data points and measures, and hence are difficult to describe. Reviewers can easily lose the forest for the trees in trying to understand such complex projects. To help reviewers in their task, grant writers should present a simple description of the basic design in the abstract and in the beginning of the method. Tables and figures can also be used effectively to clarify the study design (see Pequegnat and Stover, 1995, Appendix C).

Study Hypotheses

Clear and well-constructed study hypotheses are critical to an effective proposal because they help to establish the significance of the proposal and organize data collection and analyses (Dawes, 1995). Consequently, it is important for investigators to clarify the rationale for their hypotheses through an explicit theoretical framework and/or prior research (Herek, 1995). Figures should be used to describe the conceptual framework that is used to guide hypotheses and/or causal models that are tested.

Sampling and Recruitment of Subjects

Reviewers are interested in knowing precisely who the subjects of the proposed research are, the inclusionary and exclusionary criteria, the setting they will be recruited from, and how recruitment will be accomplished. Beyond a detailed description of subjects, reviewers look for clear documentation that will be feasible to recruit the sample size that is proposed. Feasibility can be documented by describing the numbers of potentially eligible subjects that are seen in a setting over the time period of the study. Providing evidence that similar populations have been recruited in previous research or preliminary studies in the same setting is especially convincing to reviewers.

Sample Size

Sample size is a key consideration in reviewers' evaluation of a proposal for several reasons: The costs and feasibility of a study depend heavily on sample size considerations. If a sufficient number of subjects cannot be recruited within a reasonable time period or if the projected sample size proves to be too costly to recruit, the project will not be feasible. Reviewers require assurance that the proposed project can in fact be completed. Reviewers also want documentation in power calculations that the proposed sample size is adequate to answer the major study question(s) (Cohen, 1992; Kraemer & Thiemann, 1987).

Data Collection Procedures

Reviewers respond favorably to detailed descriptions of data collection procedures, including specific measures and the time points that the data are to be

collected. It is important to include a detailed table that describes the specific data that will be collected at each time point. Reviewers want to see justification for all data collection procedures, especially for prospective studies or other intensive data collection efforts (e.g., why do measures need to be repeated; why are they repeated at the designated intervals?). Such justification is necessary to assure reviewers that additional costs of data collection are necessary from a scientific standpoint and are as cost-effective as possible.

Reviewers also want to know who will be collecting the data and how they will be trained. If a data collection procedure requires specialized training (e.g., home observations of parent-child interaction), investigators should describe the nature of the training procedures in detail. Moreover, it is important to clearly specify how necessary reliability and other quality control procedures will be established and maintained throughout the study.

Measures

Reviewers will need to understand the rationale for general and specific approaches to measurement as they try to answer the question: Why did the investigator decide to assess a particular outcome using a particular measure versus alternative outcomes and measures? Such questions are especially critical in research that involves several different outcomes (e.g., cognitive, personality, or behavioral symptoms) that can be assessed with a range of possible measures. Investigators who propose multiple measures of a complex construct should clarify their rationale for their choice and describe how data from multiple measures will be handled in the analysis.

Reviewers also want to see evidence of reliability and validity for established measures as well as a detailed plan for reliability and validation of new measures. Reviewers are favorably impressed if investigators also can provide evidence that they have used the proposed measures with similar populations in previous research and are aware of the strengths and weaknesses of such measures. Investigators who are planning to apply a measure, even a relatively well established one, to a new sample that differs in socioeconomic status, race, or culture from the measured standardization sample should provide evidence that the measure is appropriate for use in the new setting. Given space limitations, investigators may wish to put detailed descriptions and/or justification for measures in the appendix and they describe the key features of the measures in the body of the research plan.

Data Analysis

Reviewers want to see a detailed plan for each of the various phases of the proposed data analysis, for example, descriptive analyses, data reduction, measurement validation, tests of hypotheses, and additional analyses that describes how each of the measures will be analyzed (Taylor, 1995). Reviewers also want to know why the investigators chose the specific analyses that are proposed. Investigators also should clearly inform reviewers how they plan to manage common data analytic problems (e.g., type I and type II error, missing data, attrition, etc.), most especially if such problems could pose significant threats to conclusions that can be drawn from the findings.

Investigators may be tempted to propose highly technical analyses (e.g., time series analysis or hierarchical linear modeling) because they think they might impress reviewers. However, quite the opposite effect may be achieved if investigators do not provide a detailed rationale for the proposed procedures as well as evidence that they or a member of their team has the necessary expertise to conduct the proposed analyses.

Investigators also need to convince reviewers that they are conversant with the issues that are involved in the data management, including safeguards for reliability of measurement, checking of protocols, and data reduction. For this reason, investigators should provide a detailed flow sheet or diagram of the procedures and discuss special considerations that might arise in managing different types of data, for example, qualitative data, self-report, observational data, and so forth.

Human Subjects

Investigators are required to provide a detailed description of the proposed sample of human subjects including inclusionary–exclusionary criteria, anticipated number, sex, ethnic background, and health status. A detailed explanation or rationale should be provided for the involvement of special classes of subjects such as pediatric or clinical child populations, who are considered vulnerable. The rationale should clearly state why the proposed sample is critical to answer the scientific questions that are posed. Other requirements of this section include a detailed identification of the sources of research material that will be obtained from subjects (e.g., specimens, records, or data) and whether existing specimens, records, or data will be used. Investigators also should provide a detailed plan for the recruitment of subjects, consent procedures, the circumstances under which consent will be sought and obtained, who will obtain the consent, and the method of recording it.

All proposals should be approved by the appropriate institutional review boards (IRBs), ideally prior to submission but absolutely prior to the review. Failure to obtain such approvals could prevent the proposal from being reviewed. Finally, the investigator should describe any potential risks (physical, psychological, social, legal) and their likelihood and seriousness, alternative treatments and procedures that might be advantageous, and procedures for protecting against and minimizing risks including risks to participants' confidentiality (see Chapter 14, this volume). Provisions to ensure necessary interventions in the event of adverse effects to subjects and provisions for monitoring data collection to ensure subject safety should be described. Finally, the investigator should describe why the risks to subjects associated with the proposed study are judged to be acceptable in relation to anticipated benefits, including the importance of knowledge that will result from the study.

Inclusion of Women and Minority Groups

The NIH policy for research states that women and members of minority groups and their subpopulations must be included in all NIH-supported biomedical and behavioral research projects involving human subjects, unless a clear and compelling rationale shows that such inclusion is not appropriate to the health of

the subjects or the purpose of the research. To fulfill this requirement, investigators should describe how women, members of minority groups, and other subpopulations will be included in the proposed research, the composition of the proposed study population in terms of gender, racial–ethnic groups, and children and should provide a rationale for subject selection. Where minority subjects are not well represented in proposed research, investigators should say why. For some research, the lack of inclusion of ethnic minority participants may be well justified, for example, if the study group is children with cystic fibrosis, an illness that mostly affects European-Americans.

Inclusion of Children

Recently, the NIH has developed a policy to increase the participation of children in research supported by the agency (National Institutes of Health, 1998). In their research plan, investigators are now requested to provide either a description of the plans to include children and a rationale for selecting or excluding children of specific ages or an explanation of the reasons for excluding children as participants. When children are included, investigators should describe the expertise of the researchers in working with children of the proposed ages and the appropriateness of the proposed facilities for children and they should document that a sufficient number of children will be included to allow for a meaningful data analysis.

WHAT TROUBLES REVIEWERS ABOUT PROPOSALS?

Despite differences in what individual reviewers identify as problematic areas of a proposal, there are recurrent themes in their major criticisms, which West (1997) has termed “kiss-of-death” statements. The following verbatim statements taken from reviews will inform readers of the kinds of problems that reviewers often identify. For more information on this topic, readers should consult Cuca (1984) and Cuca & McLoughlin (1984).

Background–Significance

The significance of the project was not explicitly stated. It was not clear what was new or original in this proposal. The investigator did not adequately consider other relevant work in this area. It was not clear how the proposed project fits with or extends previous research.

Preliminary Studies–Previous Work

Preliminary studies are needed to justify a work of this scope. The investigator has very little evidence to support the directions of this work. How does he or she know that the proposed treatment will work? The proposed project seems premature in light of the absence of evidence to justify the investigators’ hypotheses.

Method–Procedures–Design

The design was not clear. The proposed design is not adequate to test the question; additional controls are needed. The investigator did not attend to possible confounders. The heterogeneity of the ample will make it very difficult to test the hypotheses adequately.

Subjects

Criteria for subject selection were not described adequately. The rationale

for the criteria were not specified. Recruitment procedures were not clearly described. The timetable for subject recruitment was not described.

HypothesesFramework

Hypotheses for the proposed study were not adequately described. The rationale for the hypotheses was not stated. The study does not have an adequate conceptual framework. It is not clear what is being tested in the proposed study.

Measures

Specific measures were not adequately described. It is not clear why these specific measures were chosen over alternatives. The reliability and validity of the measures were not described. The investigator has included too many measures, and their specific role is not adequately justified. The proposed measures have not been used previously with this population. It is not clear how the research team will be trained to utilize the proposed measures or establish their reliability.

Data Analytic Plan

The plan for data analysis was not well articulated. More information is needed about data reduction. The investigator intends to make many comparisons without an adequate plan for statistical correction. The power to test the hypotheses was not adequately justified. Many of the proposed analyses are not adequately justified. The investigator has included many measures that are not specified in the analytic plan. The data analytic plan was not clearly linked to hypotheses.

Feasibility Issues

The proposed project has a very broad scope. I'm not sure that it can be carried out. The investigators did not provide evidence that they will be able to obtain a sufficient number of subjects. There is no documentation of the collaborative arrangements that will be needed to implement this project as proposed.

Investigatory Team

The investigator does not have much experience or a track record of research in this area. The investigator's expertise to gather or analyze data on the measures proposed was not specified. The investigator team did not have expertise to carry out the specialized statistical analyses that are proposed. The roles of the research team and consultants were not specified.

Budget

The justification for various positions was not adequately justified. The justification for supplies, computer support, phone costs, travel support was not adequately described. The costs of consultants were not adequately supported and seem to be very high. The budget does not include sufficient time for the investigators to conduct the proposed project.

Human Subjects

The investigator did not clearly describe procedures that will be used to obtain informed consent. The ethical issues involved in administering proposed procedures were not clearly described. How the investigator will the investigator ensure confidentiality of subjects? The proposed sample underrepresents minority subjects and women.

TRAINING FOR GRANT PREPARATION

The preparation of a grant is a highly specialized, multifaceted scholarly activity that is very different from other forms of scientific writing (Baron, 1987).

Tasks such as preparation of a detailed budget and justification clearly require specialized skills as well as practice. For these reasons, didactic and experiential training concerning grant writing can be an invaluable resource that can enhance prospects for a successful grant as well as ease the burden of grant preparation. Options for training include formal courses in grant writing, mentorship, and collegial support.

Formal Courses in Grant Writing

Formal courses on grant writing can be very useful. In the Department of Psychology at Case Western Reserve University, Fagan teaches a course on grant writing that involves preparation of a research grant application in the standard format for the NIH. Course participants also present their proposals to fellow students in a mock site visit at the end of the semester. Didactic lectures on such topics as communicating ideas clearly and concisely to grant review committees who may not be familiar with the specific content area of the proposal are presented.

In developing a course on grant writing for pediatric psychologists, I have used a similar format as above but include didactic lectures concerning the nature of the review process, the components of proposals, and information about locating specific sources of funding and provide students with experience in writing different kinds of proposals, including those to foundations (see Chapter 9, this volume). We have found that it is helpful for students to serve as reviewers for their colleagues' grant proposals. Such experiences can provide students with valuable additional perspectives in the grant preparation process.

We have found that formal training in preparation of grants can help students obtain funding for their research. For example, when our research training grant was under review for a competing renewal and student stipends could not be guaranteed, we used the above seminar format to help students prepare individual research training grant applications under supervision. In this reality-based exercise, students had the experience of shepherding their applications through processing at the university's research administration department, working under the time pressure of deadlines, as well as coping with the anxieties about not getting the grant. Two of four student applicants were successful in getting their applications funded, which not only provided a career-building experience, but opened up funding opportunities for other students. Other didactic training opportunities that can be very helpful include workshops by university research administration staff that give investigators an overview of the grant application process, provide guidance in preparing revised applications.

Mentorship

Mentorship and support from researchers who have been successful in obtaining funding are perhaps the most important sources of training to help researchers to develop successful research proposals. There is no substitute for hands-on, one-on-one supervision in grant preparation from investigators who have learned hard lessons in the wars of experience as grant writers. Postdoctoral training is an ideal place to obtain training in grant writing and preparation, which can be especially

important in researchers' career development, including those in pediatric and clinical child psychology. The presence and availability of supervision and support from senior faculty mentors are invaluable to junior investigators who have the opportunity to learn the many facets of the complex process of grant applications through observations and dialogues with mentors. Moreover, junior investigators often benefit a great deal from hearing about senior faculty mentors' perspectives on the strategies involved in preparing a successful proposal. It also can be helpful for new investigators to read reviews of applications that were successful and those that were not, along with the pink sheet summaries from the review committees.

FUTURE DIRECTIONS

A critical need for the development of research in pediatric and clinical child psychology is to increase the available pool of funding that is targeted for pediatric and clinical child psychology. Achievement of this goal will require advocacy efforts that are focused on increasing the number of pediatric and clinical child psychologists on study sections at the NIH and NIMH and advocacy for research funding in areas that are critical to the development of the fields of pediatric and clinical child psychology.

Increasing the Number of Pediatric and Clinical Child Psychologists on Study Sections

Researchers in pediatric and clinical child psychology often work in cross-cutting areas that overlap with other professional disciplines, for example, child abuse. Moreover, researchers in these fields work in relatively new areas that focus on applied research, compared with the more traditional or basic areas of psychological science, that is, cognitive and social psychology. Consequently, they may face special problems in obtaining a fit between the content area and methods of their research and the expertise and experience of study section members, who may not be familiar with issues and methods in these fields. For this reason, it is important to advocate for senior scientists in the field of clinical child and pediatric psychology to be placed on study sections in agencies.

Advocacy for Research Funding in Pediatric and Clinical Child Psychology

Advocacy is needed to increase funding to be allocated for behavioral and psychological research including research related to pediatric and clinical child psychology within government agencies as well as foundations. Psychological research remains grossly underfunded compared with biomedical research, despite evidence that psychological research is important to public health (Chesney, 1997). For example, McGinnis and Foege (1993) have noted that despite the fact that behavioral factors are highly associated with the three leading sources of mortality in the United States—smoking, diet and exercise, and alcohol abuse—the national investment in prevention is estimated at only a small percent of the total annual health care cost (Chesney, 1997). Consequently, concerted efforts are needed to advocate at the congressional and agency levels (Cutlet & Cohen, 1995). We encour-

age researchers in pediatric and clinical child psychology to take every opportunity to inform local and national policymakers concerning their research programs, new findings, and relevance to public health and delivery of services to pediatric and clinical child populations.

REFERENCES

- Agnew, B. (1997). Eureka! NIH Adopts "innovation" as a grant criterion. *Journal of NIH Research*, 9, 30-31.
- American Psychological Association. (1997). NIMH restructure extramural research programs. *Psychological Science Agenda*, 10, 1-4.
- Azar, B. (1998). Peer review overhaul nears completion. *APA Monitor*, 29(8), 22.
- Baron, R. (1987). Research grants: A practical guide. In M. P. Zanna and J. M. Darley (Eds.), *The complete academic* (pp. 151-159). New York: Random House.
- Bonowski, J. G. (1996). Applying for research grants. In F. Leong & T. T. Austin (Eds.), *The psychology research handbook. A guide for graduate students and research assistants* (pp. 312-351). Thousand Oaks, CA: Sage.
- Brooks, N. (1989). Writing a grant application. In G. Parry & F. N. Watts (Eds.), *Behavioral and mental health research. A handbook of skills and method* (pp. 123-136). Hillsdale, NJ: Erlbaum.
- Chesney, M. A. (1997). NIH funding a question of behavior. *Facts of Life*, 2, 1-4.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cuca, J. M. (1984). NIH grant applications for clinical research: Reasons for poor ratings or disapproval. In *Preparing a research grant application to the National Institutes of Health. Selected articles* (pp. 453-463). Bethesda, MD: National Institutes of Health, Division of Research Grants.
- Cuca, J. M., & McLoughlin, W. J. (1984). Why clinical research grant applications fare poorly in review and how to recover. In *Preparing a research grant application to the National Institutes of Health. Selected articles* (pp. 55-58). Bethesda, MD: National Institutes of Health, Division of Research Grants.
- Cutlet, J., & Cohen, S. (1995). Influencing the political agenda on behalf of psychological research. In H. A. Pincus (Ed.), *Research funding and resource manual: Mental health and addictive disorders* (pp. 453-470). Washington, DC: American Psychiatric Press.
- Dawes, R. (1995). How do you formulate a testable exciting hypothesis? In W. Pequegnat & E. Stover (Eds.), *How to write a successful research grant application: A guide for social and behavioral scientists* (pp. 93-96). New York: Plenum Press.
- Drotar, D. (1994). Psychological research with pediatric populations: If we specialize, can we generalize? *Journal of Pediatric Psychology*, 19, 403-414.
- Drotar, D. (1995). *Consulting with pediatricians: Psychological perspectives*. New York: Plenum Press.
- Grant, A. J., & Bowe, F. G. (1995). Getting the most from federal information sources. In R. M. Eckstein (Ed.) (pp. 140-149). Coral Gables, FL: Research Grant Guides Inc.
- Grant, G. E. (1997). Memo to Direction Offices of Sponsored Programs-Office of Policy for Research Administration, NIH.
- Herek, R. (1995). Developing a theoretical framework and rationale for a research proposal. In W. Pequegnat & E. Stover (Eds.), *How to write a successful research grant application: A guide for social and behavioral scientists* (pp. 85-92). New York: Plenum Press.
- Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*. Newbury Park, CA: Sage.
- Levitant, T. (1998). (personal communication).
- Locke, L. F., Spirduso, W. W., & Silverman, S. J. (1987). *Proposals that work. A guide for planning dissertations and grant proposals*. Newbury Park, CA: Sage.
- Lorion, R. P. (1995). Grantsmanship: A view from inside and out. In W. Pequegnat & E. Stover (Eds.), *How to write a successful research grant application: A guide for social and behavioral scientists* (pp. 39-46). New York: Plenum Press.
- Lyman, W. (1995). Reading between the lines of your summary statement. In W. Pequegnat & E. Stover (Eds.), *How to write a successful research grant application: A guide for social and behavioral scientists* (pp. 73-77). New York: Plenum Press.

- Matthews, K. A. (1997) What do NIH councils do? *The Health Psychologist*, 19, 10.
- McGinnis, J. M., & Foege, W. H. (1993). Accidental causes of death in the United States. *Journal of American Medical Association*, 271, 2207–2212.
- Mucha, F. (1995). Developing a budget and financial justification. In W. Pequegnat & E. Stover (Eds.), *How to write a successful research grant application: A guide for social and behavioral scientists* (pp. 161–170). New York: Plenum Press.
- National Institutes of Health, Division of Research Grants. (1984). *Preparing a grant application to the National Institutes of Health. Selected articles*. Bethesda, MD: Author.
- National Institutes of Health. (1990). *Helpful hints on preparing a research grant application to the NIH*. Bethesda, MD: Author.
- National Institutes of Health. (1998). Guidelines on the inclusion of children as participants in research involving human subjects. *Medical Research Funding Bulletin*, April 10, 17–18.
- Ogden, T. E. (1991). *Research proposals: A guide to success*. New York: Raven Press.
- Pincus, H. A. (1995). (Ed.), *Research funding and resource manual: Mental health and addictive disorders*. Washington, DC: American Psychiatric Association.
- Taylor, H. G. (1995). Developing the data analytic plan. In W. Pequegnat & E. Stover (Eds.). *How to write a successful research grant application: A guide for social and behavioral scientists* (pp. 153–160). New York: Plenum Press.
- Walders, N., Tanielian, T. L., & Pincus, H. A. (1999). Getting funded for research. In J. Kaye, E. Silberman, and L. Pessar (Eds.), *Handbook of psychiatric education and faculty development* (pp. 190–214). Washington, DC: American Psychiatric Press.
- West, M. A. (1997). Grant-writing: A guide approval and funding. *Current Surgery*, 54, 118–123.

IV

Managing and Implementing Research with Pediatric and Child Clinical Populations

One of the primary characteristics of research in pediatric and clinical child psychology is that the data are gathered in clinical settings and research often involves complex data sets and research-related collaborations. Consequently, in order to conduct research successfully in these fields investigators need to assume significant organizational and management responsibilities such as securing and managing resources for their research, and hiring and supervising as well as managing staff. Moreover, in order to conduct research in clinical and other applied settings, researchers need to manage a number of troubling logistic problems that can be very difficult to anticipate. These include developing collaborations with agency and hospital staff, recruiting and maintaining research participants, and managing problems in data collection, especially those that threaten the integrity of study design. As one example, randomized controlled clinical trials, which are important vehicles to obtain knowledge about the efficacy of treatment, are difficult to implement and manage. In addition, investigators are increasingly called on to conduct multisite trials in order to enhance the generalizability of their research findings that have been gathered in a single site. However, these multidisciplinary trials are challenging to conduct and require specialized knowledge and methods. Finally, research with children and families in applied clinical settings raises a number of special ethical challenges concerning informed consent, confidentiality, and maintaining the role of researcher that are critical for investigators to anticipate and manage.

To address this need for researchers to obtain information about the practical problems, researchers face in designing and implementing research, the chapters in this section focus on critical aspects of managing research with pediatric and clinical child populations. Each chapter in this section also considers methods of training investigators to become more conversant with the skills and knowledge that are necessary to implement and manage research with children in applied clinical settings.

In Chapter 11, Drotar describes an investigator's eye view of relevant issues in managing research projects including anticipating tasks that are required in different phases of research, developing realistic time estimates for research tasks, working with a research team, recruiting a research team, securing time and space for one's research, managing research budgets and preparing progress reports.

In Chapter 12, Drotar and his colleagues and students consider a wide range of issues that are involved in conducting research in pediatric, school, and child welfare settings, as well as within the child mental health system. They consider strategies to anticipate and manage such key problems such as developing collaborations with staff and negotiating with them to develop and conduct research in different settings, managing communications to colleagues about research, and anticipating and dealing with obstacles to data collection and management in applied settings, including family participation and research. This chapter identifies special dilemmas that are involved in research in different settings and gives concrete examples of successful strategies in implementing research these settings.

Based on combined experience in collaborative multisite research studies, Armstrong and Drotar (Chapter 13) describe strategies and lessons from multi-institutional and multidisciplinary research trials. Using examples of multisite research focused on various populations including childhood cancer, the advantages and problems of different models of multisite collaborative research are described. The benefits of multisite research such as larger sample sizes, increased opportunities for funding and for interdisciplinary learning, coordinated statistical support and data management, and faster dissemination of research findings into practice are considered. The authors also present recommendations to anticipate and manage predictable problems in multisite studies such as maintaining standardized data collection and quality control compliance with protocols at different sites, investigator change, and difficulties in decision making. Finally, Armstrong and Drotar present models to train students to learn about and conduct multisite collaborative research.

Research with clinical child and pediatric populations raises a number of difficult ethical issues that have not been widely disseminated. In Chapter 14, Drotar, Overholser, and their students and colleagues consider a wide range of difficult ethical issues that can arise in conducting research with pediatric and clinical child populations in a range of settings. These issues include working with groups that are charged with the oversight of research ethics in specific situations, problems related to confidentiality of data, obtaining appropriate consent, managing risks related to psychological vulnerability, and maintaining appropriate role boundaries as researchers in clinical settings. Using specific illustrations from research with various populations of children and families this chapter presents a range of concrete suggestions to help researchers anticipate and prevent difficult ethical problems in conducting their research.

Each chapter in this section makes recommendations for new methods of training investigators to become more equipped with the skills and knowledge that are necessary to implement and manage research with children in applied clinical settings, for example, training in research in the context of participating in large-scale clinical trials, experiences in supervising and managing research conducted by others, and training from experienced investigators who are role models for teaching about how they manage practical problems that have emerged in their research.

11

Managing Research in Pediatric and Child Clinical Psychology

DENNIS DROTAR

In marked contrast to the image of the lone investigator conducting research in his or her laboratory, many researchers who work with pediatric and clinical child populations gather data in clinical settings and conduct research projects that involve complex data sets. Moreover, in some instances, research that advances scientific knowledge in pediatric and clinical child psychology necessitates the development of projects that involve research-related collaborations across multiple settings. To illustrate this point, consider some recent examples of studies that were published in the *Journal of Clinical Child Psychology* and the *Journal of Pediatric Psychology*. One of these reported the lessons learned from a pilot study of family-based alternatives to institution-based mental health services for youth, which involved a multiagency, collaborative effort (Henggeler et al., 1997). Another study described the results of a 4-year follow-up of the impact of home intervention on the cognitive, motor development, and behavior in play of 4-year-old children with early histories of failure to thrive (Hutcheson et al., 1997). The teams of investigators who conducted the projects assumed extraordinary organizational and management responsibilities to complete their research. Such responsibilities included (among others) securing and managing funds that were needed to conduct the research; hiring, supervising, and managing research staff; and developing collaborative relationships with investigators and professional staff in different settings.

How did these researchers develop their extraordinary management skills? What specific skills are necessary to manage research projects in pediatric and clinical child psychology? As is the case for many of the skills needed to conduct clinical research with pediatric and clinical child populations, it is very difficult

DENNIS DROTAR • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

for researchers to obtain formal training in the management of projects, especially complex, multifaceted research programs. Graduate training focuses on knowledge and skill development in research design and methods, but in my experience the skills that are necessary to facilitate the management of research are not consistently taught. Furthermore, of necessity, graduate students generally conduct research on a much smaller and less demanding scale than research conducted subsequently in their careers. For these reasons, many of the management skills that are necessary to conduct research projects need to be learned in the course of on-the-job experiences.

To address the above needs, the purpose of this chapter is to give researchers a working knowledge of the skills and tasks that are involved in managing research projects, including grant-funded research, in clinical child and pediatric psychology. Interested readers also may want to consult the description of conducting a research team and managing a lab (Molfese et al., 1996). Moreover, Chapter 12, "Conducting Research With Children And Adolescents In Clinical And Applied Settings" considers issues in implementing research. Finally, the special management issues involved in the conduct of collaborative, multi-institutional research are described in Chapter 13, this volume.

ATTRIBUTES OF AN EFFECTIVE RESEARCH MANAGER

A wide range of skills is necessary to become an effective research manager. These include interpersonal skills, leadership, problem-solving skills, and initiative.

Interpersonal Skills

Research projects with many pediatric and clinical child populations are not only time-consuming but also can involve extraordinary challenges in developing collaborations with staff in different settings (see Chapter 13, this volume). For this reason, effective execution of research with clinical child and pediatric populations requires a high level of interpersonal skills that are needed to collaborate effectively with a wide range of professional disciplines who work in different settings.

Leadership

Investigators cannot hope to carry out all the tasks that are needed to complete their research by themselves. For this reason, they need to have the ability to recruit, supervise, and motivate staff to work with them in conducting their research. Moreover, an investigator's ability to provide effective leadership of a research team is critical to the success of his or her research.

Problem-Solving Ability

The problem-solving abilities that are necessary to implement successful research projects in pediatric and child psychology also can be formidable. Investigators are called on to make difficult decisions concerning the design and conduct of

their projects, for example, subject criteria and recruitment, choice of measures, data collection, subject attrition, and so on (see Chapter 12, this volume). Moreover, given the difficulties of conducting research in applied settings, investigators are often in the position of having to anticipate obstacles that will limit implementation of their research and make informed contingency plans to facilitate the conduct of their research.

Initiative

A high level of professional initiative is needed to secure and sustain the resources that are necessary to maintain and develop research programs in pediatric and clinical child psychology. Many research projects occur in complex organizational settings such as medical schools and community agencies where investigators must compete with other projects and/or staff for space and other resources. In order to ensure that their research is implemented effectively, investigators need to become skilled in advocating for his or her staff and needs of their projects.

MANAGING RESEARCH-RELATED RESPONSIBILITIES

As described above, the abilities and skills that are necessary to become an effective investigator and manager in applied research projects are formidable. While many of these skills are difficult to learn, investigators can learn to anticipate the demands and responsibilities that are necessary to organize and lead research projects and, in some instances, obtain experience and supervision in such research. These tasks are now described.

Anticipating the Tasks That Are Required in Different Phases of Research

To be most effective in their work, investigators need to anticipate tasks that are required at different phases of their research, make realistic estimates of time for specific research tasks, define the limits of their responsibilities and resources, and allocate staff responsibilities to different tasks. One way in which investigators can accomplish their goals is by carefully anticipating the tasks that are required at different phases of their research and by allocating their time and energies accordingly. As shown in Table 1, many research projects are complex, multifaceted enterprises that involve a wide range of tasks/responsibilities, which vary as a function of the specific stages of a project. For example, investigators' tasks in the start-up phase of a project include developing job descriptions that describe responsibilities for their staff, recruiting their research team, developing collaborations that are necessary to implement their projects, conducting pilot work to refine assessment, intervention and/or recruitment procedures, and developing effective methods to implement patient recruitment.

The middle phases of a research project involve very different sets of tasks and responsibilities, such as monitoring subject recruitment, data collection, and quality control, developing methods of data management and processing, and troubleshooting concerning problems in data collection. Research projects can take

Table 1. Tasks and Responsibilities in Research Projects

Start-up phase
Develop collaborations that are necessary to conduct the research
Secure space for research
Recruit staff
Conduct pilot studies to refine assessment and intervention procedures
Develop procedures for recruitment of subjects
Secure necessary institutional sanctions for research, e.g., institutional review boards
Secure account number and administration of grant (for grant-funded research)
Develop specific procedures and manual for research protocol
Middle phase
Monitor subject recruitment
Implement quality control procedures in data collection
Develop database and data management procedures
Troubleshoot procedures, e.g., modify procedures
Prepare preliminary reports of study findings
Conduct preliminary data analysis
Clean data and maintain data integrity
Prepare preliminary publications
Prepare presentations
Grants accounting (grant-funded research)
Final phase
Complete data collection
Phase-out project staff
Consider new research projects, e.g., continuation or “spin-off” projects
Prepare publications
Prepare final reports for granting agency (grant-funded research)
Final accounting reports (grant-funded research)

many different and unexpected turns that may necessitate modification of subject recruitment procedures and study protocols (specific examples are detailed in Chapter 12, this volume). Other responsibilities in the middle phases of research include preparation of reports, presentations, and preliminary publications based on available data.

Relevant tasks in the final phases of a research project involve completion of data analysis, preparation of final reports for funding agencies, preparation of manuscripts for publications, and in some cases writing proposals for funding the next phase of a research project. In some instances, researchers also are under significant pressures to maintain funding in order to extend their ongoing research and/or to develop new projects and funding support. For this reason, they must learn to allocate time to develop the resources to support the next phases of their projects (see Chapters 9 and 10, this volume).

Developing Realistic Time Estimates for Research-Related Tasks

In my experience, it is very difficult to accurately estimate the time and difficulties that are involved in such critical research tasks as subject recruitment and data collection. As a general rule, researchers should double the initial estimates of time that they think will be needed to conduct their projects. For example, in our intervention research with children with failure to thrive recruitment and

retention of families turned out to be much more time consuming than we initially anticipated in several respects. In order to obtain a sufficient number of children and families, our team needed to collaborate with practitioners (pediatricians, nurses, and social workers) at several sites. Consequently, substantial expenditures of staff time and energy were needed to discuss the project with practitioners to ask them to refer patients to the study and to develop and maintain collaborations with professionals in different settings. Once our team secured the cooperation of referring physicians, nurses, and social workers, we faced the difficult task of identifying children who met study criteria. In order to ensure that these criteria were applied uniformly across different settings, it was necessary for our study coordinator to review hospital charts of hospitalized infants every other day at several different sites. While extremely time-consuming, this task was critical to the success of the research.

Recruitment of families also required much more time than we initially anticipated. Because families of children with failure to thrive were not consistently present during their child's hospitalization, our staff needed to spend a considerable amount of time trying to reach them by telephone and/or home visits. Finally, time-consuming efforts were spent by our intervention team and research assistants to maintain families in the project through multiple methods and to develop a network of family members who could be contacted to locate families in the event that we had difficulty sustaining contact, and so on (Drotar & Robinson, 1999). Because the time that would eventually be needed for each of these burdensome tasks was not evident to us at the outset of the study, our ability to make accurate time estimates was enhanced by pilot studies (see Chapter 22, this volume).

Because all research consumes much more time than is initially anticipated, researchers need to plan carefully for the time that will be required to conduct new projects. Such careful time estimates are especially important for grant-funded research. Investigators apply for research grants long before the funding starts but can never be certain that a proposed project will actually be funded. Consequently, when a project is eventually funded, a researcher may or may not be prepared to devote the kind of responsibility to a new project that it demands. For example, in order to manage a large-scale project effectively, investigators may need to divest themselves of some of their ongoing responsibilities, for example, teaching loads, clinical care, responsibilities, and so forth.

Allocation of Responsibilities to Multiple Professional Tasks

Investigators who conduct research projects need to develop the ability to manage multiple kinds of responsibilities at one time, (e.g., research, teaching, and/or clinical service). In order to ensure that the quality and timeliness of their research do not suffer and their work-related stress is kept within reasonable limits, investigators need to plan the allocation of their time to these multiple tasks, develop a realistic schedule for their activities, and follow it. Investigators who have difficulty in setting limits on their time and maintaining a planned schedule for allocation of time and attention to research may have difficulty completing and sustaining research programs. For these reasons, to be successful, leaders of research projects need to develop considerable self-discipline and planning ability.

WORKING WITH A RESEARCH TEAM

One of the key principles in conducting a successful research project is for investigators to develop a team of talented individuals who are given appropriate levels of responsibility for research-related tasks and supervision for their work.

Delegation of Responsibilities to Research Staff

Many successful investigators are accustomed to taking on a lot of responsibility and setting high standards for their work. Consequently, they may have difficulty sharing responsibility and turning over important and time-consuming tasks to others. To be successful in research, however, investigators have to learn to trust that others can be effective in carrying out research-related responsibilities and to support and supervise their efforts.

Deciding what and how many responsibilities to delegate to others also can prove to be a difficult task. Some responsibilities, such as data collection, need to be more clearly delegated, while others, such as decisions about implementing research in field settings, are best made by investigators in close collaboration with their staff. Moreover, investigators need to make sure that they do not delegate so many key tasks that they lose touch with their research, and hence cannot lead their project effectively.

Defining the Skills That Are Necessary for Research Positions

Another important set of skills that investigators need to learn is how to select and recruit collaborators and/or research staff who have sufficient skill, experience, and temperament to manage specific tasks that are critical to their research. Such personnel selection is best accomplished if the investigator has carefully defined the positions and skills that are necessary to conduct the research.

Defining the skills that are necessary to conduct multifaceted research projects with children and adolescents is not an easy task. In recruiting research teams in my own research, I have found it useful to distinguish among several types of key skills: (1) technical skills that are necessary to conduct research-related tasks (e.g., psychological testing, data entry, coding of tapes); (2) problem-solving skills that are needed to make necessary adjustments in methods and procedures based on experience in the field; (3) interpersonal skills that are needed to work within a research team, with collaborators, and with children and families; and (4) level of independence and initiative that is necessary to carry out the specific job responsibilities.

The skills that are best suited for specific research positions will depend on the nature of the research project and procedures. Some research tasks (e.g., coding of tapes, etc.) are discrete and highly structured, while others, such as recruitment of subjects and/or collaboration with multiple referring agencies, require extraordinary social skills and personal initiative. An example of this latter type of position was the job of project coordinator for a multisite study of the psychological adaptation of children and adolescents with hemophilia and human immunodeficiency virus (HIV) infection. The multifaceted responsibilities of this position such as facilitating management of the protocol at different sites, including maintaining

contacts with investigators, organizing data management, and troubleshooting difficulties in data collection, required a high level of problem-solving ability and interpersonal skills (Drotar, Agle, Eckl, & Thompson, 1997). Based on my experience, the skills needed to manage this position were similar to those needed to manage many of the multifaceted projects that are conducted in pediatric and applied clinical settings.

Importance of Pilot Research in Developing Job Descriptions for Research Staff

Researchers need to understand that they may not be able to define the specific tasks that are necessary to conduct their research unless they have had direct experience with data collection and project implementation. Consequently, pilot work provides critical information concerning the demands that are required by specific research-related tasks. For example, in research on psychological outcomes of infants with early history of failure to thrive, the collection of detailed home observations of mother–infant interaction involved several challenging tasks (Drotar, Eckerle Satola, Pallotta, & Wyatt, 1991), such as gathering highly detailed observational data, for example, coding in real time using a computerized event recorder and using rating scales, each of which required a considerable degree of technical skills and accuracy. However, we quickly learned that strong interpersonal skills and abilities, for example, to interact effectively with parents who were sometimes highly threatened by the observational procedures, turned out to be just as critical if not more important in implementing home observational methods as technical skills. Our observers needed to engage the families and facilitate their acceptance of the observational procedures in order to gather necessary data.

The salient emotional demands of some research-related tasks also need to be factored into the requirements for positions. For example, staff who conduct interventions for families, especially in high-risk family situations, (e.g., parents at risk for abusing and/or neglecting their children), face very difficult emotional as well as technical demands in implementing interventions in difficult circumstances. For this reason, such positions require a high level of emotional maturity, which needs to be considered in recruitment.

Issues in Recruiting Staff for Research Projects

As the above example indicates, researchers need to consider the wide range of skills, interpersonal and emotional demands, and responsibilities in developing their job descriptions and in recruiting for positions. Researchers who are under significant pressures to conduct their research with less than optimal staff resources may be tempted to use their staff in unrealistic ways and/or to create unrealistic job descriptions. Moreover, the wider range of skills that is required for a position, the more difficult it will be to identify someone to do the work effectively and the longer the search for the right person might take. Consequently, investigators may be torn between the conflicting demands to initiate their research quickly versus the need to take time with recruitment in order to obtain the “best fit” between the available talent and the demands of a research position. It takes considerable time to carefully review applications, obtain and verify reference

letters, call previous employers (which in my experience is particularly important), and interview applicants. Nevertheless, based on my experience, investigators who take the time and energy to identify and recruit individuals who are best suited for the positions in their research project will be more successful in the long run.

Evaluating Information Concerning Prospective Job Candidates

To evaluate prospective job candidates, investigators need to evaluate several different kinds of information, including the quality of the applicant's prior experience, most especially in tasks that are directly relevant to prospective research responsibilities, and recommendations, especially from other researchers. Moreover, investigators need to determine how the position fits with the applicant's professional goals, for example, the degree to which the position allows the applicant to develop skills and experience in professionally relevant areas. Research positions that fit with the applicant's personal interests and future professional goals are obviously more rewarding and motivating than those that do not.

Considering the Fit between Applicants and the Research Team

In my experience, it is especially important to carefully consider how applicants will work with colleagues who are already working on the research project. Some candidates have excellent technical skills but may have difficulty working with others on a team. Because most research teams are relatively small, interdependent groups, interpersonal conflicts on the team can disrupt the success of a research project. For this reason, the fit between the prospective job candidate's skills, interpersonal style, and interests and those of the research team is very important to consider. Moreover, wherever it is feasible to do so, the research team also should participate in interviewing and selection of applicants for research.

SUPERVISING AND MONITORING RESEARCH STAFF

Developing effective methods and structures for supervising research staff is one of an investigator's most challenging responsibilities. Questions that need to be considered by investigators in developing their methods of research supervision include the following: How much supervision is required? How often should supervisory meetings occur? How should supervision be structured? What should the purpose of supervision be?

In my experience, it is very important to develop regular and consistent forums for supervision of and support for one's research staff. The definition of "regular" and "consistent" will depend on the needs, demands, and phase of the research project. For example, very close and frequent supervision concerning the development and implementation of procedures is often critical in the early phases of a project. In subsequent phases of their research, investigators will generally reallocate their supervisory time to focus more on monitoring the quality control of data collection and management.

Individual investigators need to develop methods of supervision that work best for their own styles and specific research-related tasks. In managing my

research projects, I have generally used a combination of individual and group supervision. Individual supervision can be used effectively to focus on specific responsibilities of individual staff, expectations for their work, discussion of their personal reactions to their work, and giving support and feedback. On the other hand, group supervision can be used most effectively to facilitate a team effort to manage tasks that need to be accomplished by the entire staff or subgroup and to organize the work of the group. Depending on the size and scope of the study and staff, senior researchers may need to delegate some of the staff supervision to an associate or laboratory director or senior assistants (Molfese et al., 1996). However, investigators who delegate supervision need stay in touch with their staff and monitor the supervision.

Developing Structures for Record Keeping and Project Operation

One of the most important tasks in developing a successful project is to develop adequate structures and guidelines for research procedures and record keeping for the project. Clarity and consistency of procedures will enhance the organization and reliability of data collection and help to manage difficult issues that arise in collecting data, especially in challenging clinical settings. For these reasons, investigators should make every effort to develop written guidelines for procedures, including detailed scripts for presenting information to families concerning the project, forms for data collection, procedures for operation of the laboratory and research team, and description of modifications of procedures. It is also very useful to keep a running log of decisions that were made in implementing the research, for example, changes in administration of measures, and the reasons that they were made. See Molfese et al. (1996) for a description of sample contents of laboratory manual and procedural guide.

Developing Forums for Shared Decision Making in a Research Team

One of the characteristics of clinical research in pediatric and clinical child psychology is that research cannot always be implemented as it was initially planned and hence requires a series of difficult decisions concerning management (see Chapter 12, this volume). Consider the following example: A research team that is planning to recruit subjects from a child guidance clinic for a study of psychotherapy outcomes finds that the clinic staff is too threatened to refer patients because they feel that the researchers will evaluate their work unfairly. Consequently, decisions need to be made about how best to work with the staff so that they are less threatened by participation. If this strategy does not work, a different set of collaborators may be necessary to implement the research. Such difficult decisions often require input from the research team to obtain information concerning the progress of the research and to weigh the costs versus benefits of alternative procedures.

In my experience, the success of research in applied settings may hinge on the research team's ability to utilize their experience to make informed decisions about implementing their project. A competent research staff can provide invaluable observations and suggestions concerning changes in research procedures that are necessary to enhance feasibility and to maintain ethical standards, especially in

challenging situations. Contacts with children and families in research that is conducted in clinical settings also can raise difficult issues related to confidentiality of information and maintaining effective role boundaries that require a concerted team effort (see Chapter 14, this volume). To best use input from their staff to manage such problems, investigators need to facilitate the development of open forums in which their staff can feel comfortable in discussing problems and difficult issues (Molfese et al., 1996).

Establishing effective forums for ongoing dialogue with research staff will help to identify inconsistencies or problems in procedures that are administered by different staff and enhance standardization and quality control procedures. Such procedures are especially important to establish in large-scale, multisite projects and in studies involving complex procedures such as observations of parent-child interactions where standardization of procedures may be very difficult to achieve (see Chapter 13, this volume).

Managing Stress in the Research Team

Research, especially research with clinical populations, can engender a wide range of work-related stresses. For example, time-intensive and demanding protocols and the need to maintain children and families engaged in research can be very difficult for staff and require extraordinary persistence. Research with highly stressed families, who are difficult to contact and maintain in research, also present salient stress to staff (Drotar & Robinson, 1999). Moreover, the time-constrained demands of a research project (e.g., necessity to finish subject recruitment and data analyses by deadlines and the preparation of presentations, manuscripts, and grants), which are critical to sustaining many research projects, are inherently difficult for investigators and their staff. Finally, interpersonal stresses can be generated by differences in work styles and preferences, competition, and conflict concerning roles among members of the research team (Bernstein, 1996). Effective investigators learn to recognize signs of stress in their research team and develop strategies to help them manage them. In my experience, helping research staff to feel free to share their opinions and concerns and fostering group support can help them manage the stress.

DEVELOPING DATA MANAGEMENT AND ANALYTIC CAPABILITIES

The development of capacities for data management and analysis are critical to effective monitoring of quality control and rapid production of progress reports and manuscripts, and hence are important in sustaining a successful research program. Critical tasks involve setting up a database and procedure for monitoring data entry and tracking accuracy of data entry.

Moreover, many of the projects that are conducted in pediatric and clinical child psychology involve large data sets, for example, multiple measures that are repeated over time to relatively large numbers of research participants. Such data sets are very demanding from a data management standpoint and require specialized expertise from research staff and specialized data analytic systems to facilitate accurate data entry and management (Thompson, 1994).

In my experience, it is very important to develop databases and systems for tracking data collection and data management as soon as possible. This will allow more efficient data reduction and more rapid assessment of the performance of key instruments and procedures. To ensure effective quality control of their data, investigators need to develop effective methods to monitor the pace of data collection, such as windows in which various procedures should be completed, as well as the quality of the data that are being collected. Unless the investigator and research team closely monitor the pace and quality of data collection, they will not be in a position to identify potential departures from the research protocol as soon as they occur and to institute corrective action. Consequently, data may be lost or rendered useless.

Another important reason to establish capacities for data entry and management very early in the course of the research is to establish procedures for data reduction that are necessary for an optimal data analytic plan. To the extent that investigators carefully plan their data reduction, they will also implement a more focused data analytic plan that will reduce unnecessary analyses and statistical tests. Careful data reduction is also advantageous to conduct necessary measurement studies. For example, investigators who develop measures for purposes of their study need to refine them through more extensive data analysis. Moreover, if measures are applied to populations other than that from which they were originally derived or standardized, investigators need to develop a plan to assess the validity of such applications (see Chapter 5, this volume).

SECURING AND MAINTAINING RESOURCES TO CONDUCT RESEARCH

Securing and Maintaining Time to Direct the Project

Because time and commitment are critical resources in the eventual success of a research project, investigators should include a sufficient degree of funding in their research proposals to allow them to participate in the research and ensure that the time they have allocated for their research matches what is budgeted. Even when sufficient funding is available to cover investigators' time commitments to their research, it may be difficult for them to reallocate their job responsibilities, especially in clinical settings. Moreover, colleagues may be frustrated and disappointed by investigators who need to take time from clinical and teaching activities in order to conduct their research. For example, when I made the transition from nearly full-time clinical work to directing a research project concerning clinical intervention for families of children who were failing to thrive, which took up more than half of my time, most of my pediatric colleagues were upset because they knew that the research would limit my clinical work.

As this example indicates, investigators need to consider the implications of change in job responsibilities when they begin new research projects. For example, if one needs to reduce one's clinical or teaching time in order to conduct a new research project, is there someone who can cover these activities? What will be the implication for relations with colleagues?

Negotiating Time for Research

The level of support from one's administrative director for research activity is a critical factor in securing and maintaining time for research. Administrative support is especially important to help investigators develop realistic allocation of their job responsibilities to conduct their research. In some settings, institutional administrative policies can limit optimal time allocation for research. For example, administrators may regard new research grants as solely a means of additional fiscal support. In fact, it is in the best financial interest of the institution to have the investigator take the salary for a new grant and continue many of their other responsibilities as they were doing previously. While this "solution" may fit institutional aims, it may be quite untenable from the investigator's and granting agency's perspectives. Consequently, it is important for investigators to forcefully negotiate with their administrators to ensure that they have sufficient time to conduct their research and to have some assurance that their time that is supported through research grants will be reasonably protected from other responsibilities. Although these negotiations are difficult to manage, investigators need to be persistent to ensure that they succeed at their research.

Negotiating for Research Space

Next to negotiating changes in the allocation of time and job responsibilities to conduct research, space is clearly the most contentious of resource-related negotiations. Many applied settings, especially hospitals, are bursting at the seams with staff. Consequently, there is often a high level of competition for research space for obvious reasons. If research staff do not have sufficient space to see children and parents and carry out necessary procedures, it may be impossible for them to carry out their projects effectively. Moreover, a lack of adequate space can have a detrimental impact on research staff's morale, and hence limit their productivity.

Given the importance of research space and the difficulty of these negotiations, it is important that investigators advocate early, persistently, and at times vociferously for necessary space. Unfortunately, many institutions do not have clear, rational, or consistent policies for the allocation of research space. Consequently, investigators should advocate forcefully for their research needs based on their project needs, the level of their funding, and the potential for future funding.

Another important issue in negotiating for space, especially in medical settings, is that the unique needs for space that are required by psychological research generally are not apparent to those who are responsible for making decisions about space allocation. For example, some psychological research projects may require a great deal of space, for example, to conduct follow-up assessments, to see children and families for developmental testing, and to provide videotaped assessment of parent-child and family interaction. Consequently, investigators need to educate research administrators in their settings concerning the specific activities and procedures in their research and the space that is required to complete them.

MANAGING RESEARCH BUDGETS

Investigators who conduct grant-funded projects are responsible for the management of their projects' budgets, which is a critical task that they may need to

assume but for which they have little training. Many universities and hospitals have grants management and budgeting offices to facilitate this task and investigators should avail themselves of as much help and training from these sources as they can. Moreover, many settings, especially those that manage a large number of research grants, have administrative staff who provide fiscal grants.

On the other hand, in my experience, even if their setting has excellent grant management capabilities, investigators should not rely solely on institutional resources for budget management and should take the initiative to develop their own systems to manage and track their expenditures. Even investigators who have been careful in developing their budgets may underestimate or overestimate various research expenses. However, careful monitoring of project expenses is especially critical, because many granting agencies generally require investigators to assume significant cuts in their proposed budgets once their projects are funded. Consequently, the more closely that investigators monitor expenditures of their project, take steps to limit expenses that are out of line with their projections, and cut costs wherever they can, the more likely that troubling cost overruns will be prevented. In the event of overruns, government agencies will not fund any more than the award. Moreover, an investigator's institution cannot be counted on to make up for this funding.

To effectively monitor research-related expenses, a project's internal record of expenditures should be checked against the accounting statements to monitor progress spending and to resolve inconsistencies. In my experience, it is not uncommon for incorrect charges from various sources to be billed against one's grant. For this reason, savvy investigators will carefully look over their expenses in detail to determine whether these expenses were billed correctly and institute corrective action to have the charges removed.

PREPARING PROGRESS REPORTS

Successful investigators manage the timely completion of reports concerning their project's progress to relevant agencies. Typical reports describe the overall progress of a project, accomplishments in data collection and analysis, and problems; include manuscripts; and involve a detailed budgetary report. While investigators can and should rely on their staff to assist in the preparation of research-related reports, it is important that they assume the primary responsibility for ensuring that reports for funding agencies are completed in a timely and thorough manner. Failure to do so could jeopardize current or future research funding.

In some instances, progress reports will be used in grant renewals. Review committees weigh these progress reports heavily in evaluating a grant (see Chapter 10, this volume). Consequently, investigators should take the requirements and deadlines for these reports seriously and document changes in the protocol, including the reason changes were instituted, analyses that were conducted, and so forth.

TRAINING FOR MANAGEMENT OF RESEARCH PROJECTS

One important gap in educating researchers in pediatric and clinical child psychology concerns training in management of research projects. It is ordinarily

very difficult for graduate students to receive training in the management of large-scale research projects. Nevertheless, we have developed several ways of enhancing students' management skills and opportunities for training in management of research with pediatric and clinical child populations. These are described in the next section.

Didactic Training

Investigators who have had extensive experience in managing research projects are in a unique position to provide didactic training to help students understand and master critical skills (e.g., management of staff, data analytic management, ethical issues in managing research). One vehicle for such training is a lecture series in which experienced investigators discuss various facets of their research programs including management of the research. Many of these investigators have conducted complex research projects such as prospective follow-up of very-low-birth-weight infants or cocaine-exposed infants have presented substantial challenges for management and implementation. Our students have gained a great deal from their dialogues with investigators concerning how they managed difficult logistical issues, for example, methods of recruiting participants and reducing subject attrition.

Supervised Training in Managing Research

There is no substitute for supervised experiential training in managing complex, multifaceted research projects. Several types of experience can be most helpful, including job experience in large-scale projects in clinical settings, supervised experience as part of a research team, and managing research in applied settings.

Practical Experiences in Managing Research

Practical experiences in managing and conducting research prior to graduate school provide an excellent foundation for students to learn how to implement research as well as other skills (e.g., data collection and analysis, writing protocols for the institutional review boards, grants management). Students who have had experience in coordinating and developing research programs prior to graduate school generally have utilized these skills to great advantage by successfully managing difficult research projects in applied settings for their master's and dissertation research. For this reason, I would encourage undergraduate students who are interested in graduate training in child clinical and pediatric psychology to obtain such experiences.

Supervised Experience as Part of a Research Team

Students also can gain invaluable experiences in managing research by conducting their research as part of ongoing, large-scale projects that are organized and supervised by faculty. Several of our students have conducted their clerkship,

master's, and/or dissertation research with data from ongoing prospective studies [e.g., prospective follow-up of the impact of head injury on children and families (Taylor et al., 1995), developmental outcomes of Ugandan infants with HIV infection (Drotar et al., 1997)]. In this work, students have become involved in the activities of these projects, have observed and participated in assessment procedures and planning meetings, have worked with project staff to implement the protocols, and eventually have developed their own research data set. Participation in such projects has given our students excellent opportunities to observe the workings of a successful research project, including organization of data collection and methods of analysis.

Another way that students can obtain excellent experience in management of research projects is to join a laboratory group that includes students and staff in projects that are directed and supervised by their mentors (Molfese et al., 1996). By participating in such projects, students then can gain experience in different roles and responsibilities that are commensurate with their level of training. For example, students may initially work in more restricted roles, for example, in gathering data, and subsequently assume leadership roles such as supervising and monitoring the work of research assistants, grant writing, and/or management.

Managing Research in Applied Settings for Masters and Dissertation Research

Students also can gain valuable experience in managing research, including organizing and supervising the work of others, or through their own research, especially their dissertation research. Many students in our graduate training program have developed dissertation research projects in which the scope of data collection has necessitated the efforts of a number of other students to conduct data collection and entry. Consequently, our students not only have gathered their own data, but also supervised data collection and interreliability studies conducted by other undergraduate and graduate students.

To maximize the learning opportunities that are involved in such work, it is important that students receive mentorship concerning the supervision of others (e.g., the need to set specific goals, the need to maintain accountability in work of research assistants). To reap maximal professional benefits from their experiences in conducting their own research in applied settings, students benefit from the close mentorship of senior researchers who are experienced with such research and hence can help their students avoid some of the pitfalls (see Chapter 12, this volume).

Postdoctoral Research Training

Postdoctoral research training experiences provide an excellent opportunity for investigators to learn skills in the management and implementation of research under the tutelage of experienced mentors. Postdoctoral research training programs can provide direct supervised experience for fellows to develop and manage their own research projects and supervise staff, experience in grant writing, and so forth, under the supervision of experienced mentors.

Continuing Education Opportunities

Researchers who are involved in managing complex research projects involving children and families should avail themselves of courses and didactic experiences concerning such topics as grants management, supervision, managing difficult issues, for example, conflict resolution, data analytic approaches that are appropriate for complex data sets, and so on. Moreover, investigators can benefit from learning how researchers have dealt with difficult issues in the management of their research. For this reason, I would also encourage the development of symposia and roundtable discussions at professional meetings to focus on issues and strategies in implementing clinical research with children, adolescents, and their families in a range of settings. Moreover, written descriptions of the lessons learned by senior investigators in managing and implementing research with different populations and different settings would be highly instructive (for an example, see Spiker, Kraemer, Scott, & Gross, 1991).

REFERENCES

- Bernstein, E. (1996). Diversity in work styles. In F. T. L. Leong & J. T. Austin (Eds.). *The psychology research handbook. A guide for graduate students and research assistants* (pp. 325–349). Thousand Oaks, CA: Sage.
- Drotar, D., Agle, D. P., Eckl, C. L., & Thompson, P. A. (1997). Correlates of psychological distress among mothers of children and adolescents with hemophilia and HIV infection. *Journal of Pediatric Psychology*, 22, 1–14.
- Drotar, D., Eckerle, D., Satola, J., Pallotta, J., & Wyatt, B. (1990). Maternal interactional behavior with non-organic failure to thrive infants: A case comparison study. *Child Abuse and Neglect*, 14, 41–51.
- Drotar, D., Olness, K., Guay, L., Marum, L., Wiznitzer, M., Hom, D., Svilan, M. S., & Fagan, J. F. (1997). Neurodevelopmental outcomes of Ugandan infants with human immunodeficiency virus type I infection. *Pediatrics*, 100, E1–E5.
- Drotar, D., & Robinson, J. (1999). Researching failure to thrive: Progress, problems, and recommendations. In D. Kessler & P. Dawson (Eds.), *Failure to thrive in infants and children: A transdisciplinary approach to nutritional adequacy in children* (pp. 77–98). Baltimore: Paul H. Brookes.
- Henggeler, S. W., Rowland, M. D., Pickrel, S. G., Miller, S. L., Cunningham, P. B., Santes, A. B., Schoenwald, S. K., Randall, J., & Edwards, J. E. (1997). Investigating family-based alternatives for institution-based mental health services for youth: Lessons learned from the pilot study of a randomized field trials. *Journal of Pediatric Psychology*, 26, 226–233.
- Hutcheson, J. J., Black, M. M., Talley, M., Dubowitz, H., Berenson-Howard, M., Starr, R. H. Jr., & Thompson, R. S. (1997). Risk studies and home intervention among children with failure to thrive: Follow-up at age 4. *Journal of Pediatric Psychology*, 22, 651–668.
- Molfese, D. E., Murray, K. L., Martin, T. B., Pesers, C. J., Tan, A. A., Gili, C. A., & Simus, P. A. (1996). Coordinating a research team: Maintaining and developing a good working laboratory. In F. T. L. Leong & J. T. Austin (Eds.), *The psychology research handbook. A guide for graduate students and research assistants* (pp. 311–324). Thousand Oaks, CA: Sage.
- Spiker, D., Kraemer, H. C., Scott, D. T., & Gross, R. T. (1991). Design issues in a randomized clinical trial of behavioral intervention: Insights from the Infant Health and Development Program. *Journal of Developmental and Behavioral Pediatrics*, 12, 386–393.
- Thompson, P. (1994). Data warehouse manager. Cleveland, OH: Department of Radiology, Case Western Reserve University School of Medicine.
- Taylor, H. G., Drotar, D., Wade, S., Yeates, V., Stancin, T., & Klein, S. (1995). Recovery from traumatic brain injury in children. The importance of the family. In S. Broman & M. E. Michel (Eds.), *Traumatic head injury in children* (pp. 188–216). New York: Oxford University Press.

12

Conducting Research with Children and Adolescents in Clinical and Applied Settings

Practical Lessons from the Field

**DENNIS DROTAR, JANE TIMMONS-MITCHELL,
LAURA L. WILLIAMS, TONYA M. PALERMO,
RACHEL LEVI, JANE R. ROBINSON,
KRISTIN A. RIEKERT, and NATALIE WALDERS**

In order to successfully conduct research in clinical and other applied settings with children and adolescents, investigators need to learn to manage a number of logistic problems that can be difficult to anticipate (Drotar, 1989). These problems include developing collaborations with agency and hospital staff that are necessary to recruit subjects (Drotar, 1993), recruiting and maintaining research participants in studies, and managing problems in data collection, especially those that threaten the integrity of study design. Researchers who work with children and families need to anticipate as many of these problems as possible so that they can either implement strategies to prevent them, which is the preferable approach, or develop data analytic approaches to limit their influence on the quality of their data (see Chapter 4, this volume).

DENNIS DROTAR, TONYA M. PALERMO, and JANE R. ROBINSON • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038. **JANE TIMMONS-MITCHELL** • Department of Psychiatry, University Hospitals, Cleveland, Ohio 44106. **LAURA L. WILLIAMS** • Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, South Carolina 29425. **RACHEL LEVI, KRISTIN A. RIEKERT, and NATALIE WALDERS** • Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123.

In our experience, it is difficult for researchers to anticipate and manage the specific problems that are involved in conducting research on children and adolescents in applied settings unless they have experienced them directly. Moreover, information about these problems is generally not published or widely disseminated in scientific presentations. Consequently, many practical problems that are involved in implementing research projects in clinical settings and the strategies that are most useful to circumvent them remain well-kept secrets, especially from novice researchers.

To address the need for better description of the common practical problems that arise in conducting research in clinical settings with children, the present chapter summarizes some of the issues and obstacles that have been encountered in conducting research with children in a number of settings such as pediatric hospitals, schools, child welfare, and mental health agencies. Based on experiences in conducting research, this chapter summarizes these problems as well as strategies that have been found to be useful in managing them in different settings.

Researchers who are interested in learning more about process-related issues in conducting research with children and families may wish to consult Brody and Sigel's (1990) volume, which includes researchers' biographies of research projects with clinical populations of children and families. Moreover, Chapters 21 and 22, this volume, consider methodological and practical issues involved in evaluating mental health services with children and families.

RESEARCH IN PEDIATRIC SETTINGS

The conduct of research in pediatric hospital settings raises several challenging issues such as working with other professionals, obtaining resources to conduct research, and managing obstacles to data collection and barriers to family participation, some of which also are encountered in other settings with children.

Ways to Improve Collaboration with Pediatricians and Other Professionals Concerning Research

One of the primary characteristics of research with children in any clinical setting is the need to rely on service providers to identify eligible research participants and facilitate their recruitment. This is most certainly the case in pediatric and other medical settings where the cooperation of the child's physician is absolutely essential to implementing research. Such research collaborations do not occur spontaneously; rather they need to be carefully nurtured and developed. Readers are referred to Chapter 12, this volume, and Drotar (1995) for descriptions of research-related collaborations with pediatricians. Moreover, Chapters 6 and 24, this volume, each provide excellent descriptions of programmatic collaborative research with pediatric subspecialists.

Securing Time and Resources to Support Data Collection

One of the key issues in developing research in any applied setting, especially medical settings, involves developing the necessary resources to support the time

involved in data collection. In our experience, it is simply unrealistic to expect colleagues who are heavily involved in service delivery to devote very much time to data collection unless they are compensated to do so. Moreover, even with funding, this is difficult to accomplish. Consequently, researchers in pediatric (and other) settings need to develop resources to support data collection through research funding (see Chapters 9 and 10, this volume) and/or by obtaining help from community and student volunteers.

Involving Practitioner Colleagues in Research

Researchers who work with practitioners in clinical settings, including pediatricians, should not assume that the goals and purpose of their research are necessarily important or clear to their colleagues. We have found that multiple forms of communication (e.g., brief oral presentations and written summaries of research plans) are particularly useful in informing colleagues about the nature and importance of their research. Wherever it is realistic to do so, the potential clinical relevance of psychological research to assessment and intervention should be emphasized. The more that pediatric staff believe that psychological research is potentially relevant to clinical care, the more likely they will want to help with this work. In addition, the more that pediatricians and support staff can participate in developing a research project, the more they will support the study. Consequently, wherever it is feasible to do so, involving colleagues early in the research planning process is clearly advantageous.

Psychologist researchers who work with pediatric populations also should recognize that it is in their best interest to work closely with all professional staff who have contact with pediatric populations in implementing research in a particular setting. For example, nurses and social workers have long-standing relationships and extensive contacts with children and families in many pediatric settings that are critical to implementing psychological research in medical and other service settings (Drotar, 1989).

To encourage professional staff's interest in research, psychologist researchers should spend time getting to know them and their work, show an interest in them, and acknowledge their help with the research. Concrete acknowledgments such as thank you notes, as well as occasional food and beverages, are in our experience very much appreciated by colleagues in clinical settings.

Special Obstacles to Conducting Research in Pediatric Settings

Researchers need to anticipate several frequently encountered obstacles to conducting research in settings. These include colleagues' communication to research participants, practical problems in data collection, and barriers to family participation. With sufficient recognition and planning, we have found that some of these problems can be anticipated and prevented.

Colleagues' Communications to Research Participants

Medical colleagues may sometimes unwittingly behave in ways that can threaten the quality and/or ethics of data collection. For example, a graduate stu-

dent developed her dissertation research in close collaboration with a pediatrician whose eagerness to help with the research created problems. In an effort to recruit subjects, he offered incentives to families that were not available to them, for example, copies of videotapes for families. Moreover, he told colleagues that they would receive feedback concerning observations of parent-child interactions that were research data and confidential to the investigators. Fortunately, the good working relationship between the student and the pediatrician allowed open discussion and eventual agreement on the pediatrician's part to refrain from making unwarranted statements about the research to families and colleagues.

Physician collaborators are often very interested in helping to facilitate subject recruitment, but sometimes do not consider the methodological problems that can arise in this process (see Chapter 4, this volume). For example, we have found some practitioners have their own criteria for the definition of clinical problems such as failure to thrive, which do not always correspond to more stringent research-based definitions (Drotar & Robinson, 1999). Moreover, in an attempt to be "helpful" to investigators, some practitioners may screen out parents with severe psychopathology or identify only those families who they believe are the most motivated for research.

To help manage and prevent such problems, we have found that it is useful to inform pediatric colleagues about the research design, their potential role in recruiting families, the specific criteria for sample selection, and to emphasize the critical importance of applying these criteria uniformly to reduce sampling bias. Developing standardized scripts for communicating with families about their participation in research is another very useful strategy.

Practical Problems in Data Collection

Data collection in pediatric settings poses several practical challenges. One of these is very basic: that is, where to collect the data. For some projects (e.g., those in which the data collection is restricted to questionnaires or interviews), collecting data in clinic settings may be optimal, given the advantages of visibility to pediatric colleagues and convenience for children and families. On the other hand, data collection in clinic settings can pose additional burdens on staff logistical (e.g., competition for space, problems in keeping track of families, etc.), and ethical problems, such as confidential collection of data (see Chapter 14, this volume).

In clinics where the quality of data collection is significantly limited by such problems, investigators should consider other options for data collection, such as home or laboratory settings. We have used home-based data collection in studies that involved time-consuming interviews to document stressors and coping strategies among children with cancer. These would have been very difficult to conduct in a clinic setting (Bull & Drotar, 1991). On the other hand, data collection in the home requires substantial staff resources, may raise family concerns about privacy, and may not always be feasible, for example, for families who live far from the center. In order to maintain standardized procedures, data collection in a research laboratory or office is ideal. Unfortunately, such space is not available in many clinical settings.

Barriers to Family Participation and Completion of Research Protocols

One of the most challenging issues in data collection in pediatric and other applied clinical settings is how to enhance family participation in research. Families of children with various clinical problems may be highly stressed by financial and family problems, the demands of child care, and problems with access to transportation that limit their availability and readiness to participate and complete research protocols. Such parents can be difficult to contact to discuss their participation in research, let alone secure their participation. Others may feel threatened by their potential involvement in psychological research, especially if they feel that their parenting or family life will be criticized. Parents from ethnic minority groups may have special concerns about being harmed or scrutinized by their participation in research (see Chapter 8, this volume, for a discussion of research with ethnic and minority children).

Strategies to Enhance Data Collection

Strategies that we have found to be helpful to facilitate data collection and the management of research in pediatric and other medical settings include understanding the operations of the setting and staff, becoming useful to the medical team, informing the staff about research progress, making changes in procedures to facilitate data collection, and developing strategies to enhance family participation in and completion of research.

Understand the Operations of Medical Settings and Staff

Each medical setting operates somewhat differently. For this reason, investigators should identify the key staff in the specific setting from which their sample will be recruited and learn their preferences concerning how their research can be best conducted. A comprehensive understanding of clinic operations and the roles and responsibilities of various staff will help investigators to fit their research into the setting without making undue demands on staff or disrupting clinic procedures. For example, to help prepare for collection of data in a study of quality of life of children with cancer and their parents, one investigator spent the day shadowing a nurse in the oncology clinic. The time spent in such observation and conversation helped to increase the investigator's comfort and the staff's awareness of her research.

Become Useful to the Medical Team

In order to gain the staff's cooperation with research protocols, we found that it is important to be as helpful to staff in their work as is possible. For example, in a study of efficacy of preparation for children for surgery, the investigator was not notified when the patients had arrived as had been agreed upon. To address this problem, she offered to transport the patients from the admitting department to surgery as well as study patients. This offer was helpful to staff and saved everyone some time. After offering this service, she was notified promptly each time a potential research subject arrived.

Answer Questions and Inform Staff of Research Progress

In order to communicate the value and importance of psychological research to staff in medical settings, it is very important to keep them informed of research progress and to give them an opportunity to ask questions and/or register their concerns about how the work is progressing. Once the study is completed, it is critical to communicate the study results to the staff and provide an opportunity for them to ask questions and discuss the clinical implications of the research. We have found that pediatric practitioners greatly appreciate the opportunity to hear about the results of psychological studies concerning the children and families in their practice.

Make Procedural Changes to Facilitate Data Collection

In order to facilitate their data collection, investigators who work with pediatric and clinical child populations often need to modify procedures that were initially planned. Such changes need to be negotiated with staff, most especially if they require them to change their customary methods of operation. For example, in collecting data concerning children's coping and expectations of surgical procedures, an investigator found that she had a very small window of opportunity to gather data from families while they were in the preadmission testing visit. If she waited until the families finished with their visits, she would have had to spend an additional 20–25 minutes in the hospital. For this reason, she needed to develop an alternative procedure. The strategy that proved to be most effective was to break the protocol down into several parts: Informed consent and demographic forms were filled out before the children were seen by the nurse and surgeon. The other study questionnaires were then fit in between the nurse and the surgeon visits.

Other changes in research protocols in pediatric and other applied settings are necessary when the researcher discovers (ideally early in the research process) that his or her initial plans for data collection were unrealistic. For example, a student initially utilized three ambulatory care clinics to collect data for her master's research, which focused on adolescents with asthma. Pilot work in these settings indicated that data collection proved to be much slower than anticipated. In order to secure a sufficient sample size, inpatients were also included, which required establishing relationships with a new set of staff and developing new procedures for data collection.

Investigators in hospital settings also may encounter problems in staff cooperation concerning data collection that necessitate changes in their procedures. For example, in a study of children's responses to pain following inpatient surgery, an investigator encountered marked variation in nurses' compliance in filling out pain ratings, which was an important dependent measure. In order to enhance the staff's cooperation, the investigator came to their meetings to discuss the project; moreover, the staff was asked for their input about how the data could be collected more effectively. Following an in-service for the nurses concerning collection of pain rating procedures, other adjustments were made in the study's procedures, which proved to be successful: For example, to facilitate collection of pain ratings from anesthesiologists, the investigators held up the response sheets in front of them and

asked for their answers while marking them down. Reinforcements in the form of a charm pop for each pain rating collected were also given to the nurses.

Develop Strategies to Enhance Family Participation in and Completion of Research

To limit bias due to selective participation and attrition, which is an important methodological problem in pediatric and clinical child psychology (see Chapter 4, this volume), researchers who collect data in clinical settings need to implement procedures to enhance family participation in research. We have identified several methods that have been successful in improving family participation rates in pediatric and other clinical settings.

Increase the Convenience of Data Collection for Families. In order to implement a data collection plan that fits with both family schedules and clinic routines, investigators must become very familiar with the routines of patients in particular settings. One effective strategy is to make sure that the data collection procedures are tailored to the family's schedule and clinic routines, for example, arranging the data collection to occur within periods of time where patients are already waiting for their medical appointments. Investigators who have sufficient staff resources also can offer flexible scheduling arrangements such as weekends and evenings for families who are interested. Moreover, a high level of investigator flexibility is necessary to collect data from some groups of research participants, such as fathers.

Present the Study to Families Clearly and Positively. In our experience, psychologist researchers who collect data in clinical settings should present the basic information about the study themselves rather than rely on others to do so. Researchers are the most knowledgeable about their studies, and hence are in the best position to present them to families for their consent. On the other hand, to preserve the family's confidentiality, the first contact to inform them about the study by letter or in person is generally made by the practitioner, for example, pediatrician, who has the continuous professional relationship with child and family. For this reason, practitioners' support of and enthusiasm for the study are necessary to enhance the level of family participation.

We have found that it is generally useful to present the study to children and families in a positive, upbeat manner and to describe what will be required from those who participate. In order to be most credible with children and families, investigators certainly should be well-versed in their methods and present a convincing rationale to help families understand why it is important for them to participate.

Investigators also should ensure that they spend a sufficient amount of time with families to carefully discuss the purpose of the study and the nature of their participation. Ample opportunity for dialogue between families and investigators will help to ensure that families understand what is expected of them in the research (see Chapter 14, this volume). Moreover, it is important to make sure that the discussion of participation and informed consent takes place at a time that is

convenient for families and allows a nonhurried, respectful presentation of the study.

Finally, we also have found that it is useful to not accept a "no" response from a parent without first discussing and clarifying families' concerns. For example, in discussing their concerns about participating in a study, a family may reveal that they were most concerned about a relatively small phrase in the informed consent procedure that they may have misinterpreted. In some instances, when such concerns are clarified and parents understand what is expected, they may indicate an interest in participating. On the other hand, researchers who have such dialogues with families must be careful that they are not in any way coercive or perceived as such by families.

Reduce Practical Barriers to Family Participation. In order to maximize participation in research and reduce attrition, especially in prospective studies, researchers need to understand and address practical barriers that may limit families' participation and develop ways to increase families' access to research. For example, increasing parents' access to transportation by providing reimbursement for bus fare, coupons for taxis, and payment for parking may increase family participation in research (Drotar & Robinson, 1999). Similarly, in some studies it may be necessary to provide supervision for siblings in order to relieve parents of the burden of finding child care while they participate in research.

Successful data collection in clinical settings often requires a great deal of ingenuity and persistence. For example, during a study of mothers' response to practitioners' referrals for psychological intervention for depression in a pediatric primary care setting, we encountered many problems (i.e., disconnected phones and wrong numbers; subjects without a working phone who relied on friends and family members for phone access; pager numbers instead of operating phones, etc.). To manage these problems, hospital databases were checked frequently and charts were reviewed to determine if mothers left new phone numbers when leaving messages for doctors.

Several steps also can be taken prior to scheduled appointments to maximize families' and children's participation rates. For example, once the research appointment is scheduled, it may be helpful to send a letter with the pertinent information and/or reconfirm with a phone call one day prior to the appointment.

Provide Incentives to Families to Participate in Research. Whenever it is feasible to do so, investigators should provide incentives for families to participate in research. For many families, a concrete incentive can make an important difference in their participation. Offering families payment for their time or other incentives such as copies of videotaped interactions of their child can enhance their participation. In addition, children often enjoy receiving a small "prize" or toy for participating in a study.

Stay in Close Contact with Research Participants. When working with families who have agreed to participate in prospective studies, it is important to maintain close contact over the course of the research in many ways as are feasible given the investigator's resources. For example, a thank you letter not only provides

an expression of appreciation for a family's participation in research, but provides a check on whether their addresses are up-to-date. Moreover, giving families a staff member's phone number should they have questions sends a clear message about the investigator's interest in the family. Birthday cards for the child and mother, holiday greetings, and/or regular phone contacts to inquire about the child's progress are other helpful ways to keep families aware of the study and motivated to participate (Cicchetti & Manley, 1990).

To help maintain family members' interest in the study, it is also helpful to send out progress updates that summarize findings, results, and implications (Drotar & Robinson, 1999). To be most effective in maintaining contact with parents, we have found it helpful to obtain several phone numbers with additional addresses. Parents are usually willing to give a phone number and address where messages can be left (e.g., their mothers, sisters, or friends), which is very important given the frequency of moves and phone disconnections in some populations. Detailed information about the extended family network also can be useful in maintaining contact with families.

ISSUES IN IMPLEMENTING RESEARCH IN SCHOOLS

Research with children in schools raises some special advantages as well as problems (Bond & Compas, 1989; Meyer, Miller, & Herman, 1993). One important benefit of collecting data within a school setting, especially a public school, is access to a large, diverse group of students. Relatively brief surveys also can be administered to a large group of students simultaneously, thus easing the burden of data collection for the researcher. Despite these advantages, researchers need to manage a number of obstacles in order to implement research in schools such as convincing staff of the value of research, securing cooperation from teachers, and enhancing student and parent participation in research.

Convincing Staff of the Value of Research

As is true for research in health care settings with children, probably one of the most important issues in conducting research in school settings is the need to establish working relationships with key individuals such as teachers and key administrators, especially the superintendent. While some school systems, especially those that are located close to universities, have established protocols for handling research requests, many smaller districts do not. For those settings without established procedures, researchers must take primary responsibility for educating key individuals about the importance of the proposed research project and its logistics. Because some school system administrators and/or board members are unfamiliar with the research process, gaining their trust through open dialogue about the goals of the project, instruments to be used, and so forth, is critical to the success of research in schools.

As is the case for research in hospital settings, identifying the practical relevance of the proposed research and carefully explaining how the research may eventually help the school system to better serve their students, for example, how information gained from research may help to design interventions, is a useful

strategy. For example, suppose a researcher wants to gather data concerning the rates of exposure to and victimization from violence among middle-school students and the coping strategies that are utilized by early adolescents to manage such stressors. School administrators can be informed that they might be able to use the data from such a project to organize support groups for students who have been exposed to violence and to heighten faculty awareness about the impact of violence exposure and sensitivity toward students' needs.

Once a good working relationship has been established with system-level administrators and approval has been gained, access to schools and teachers is often much easier. Nevertheless, the fact that this process may take multiple contacts with multiple staff should be considered in planning for research.

Managing Requested Changes in Procedures

Another point worth noting is that even when investigators are granted access to a school for data collection, significant constraints can still be placed on the kind of information that can be gathered in psychological research. For example, questions about students' sexual behavior or drug use may not be permitted in some schools. While administrators' primary goal in limiting such inquiries is to protect students and their school system, the quality of data that may be obtained as a consequence may be skewed and/or seriously limited. Consequently, researchers need to be aware that the instruments and questions that they initially presented for the school staff's consideration may not be the versions that are approved.

In cases where procedures are changed, investigators may have to make a very difficult decision about whether or not it is fruitful to proceed. Essentially, the investigator must ask him- or herself if his or her ability to examine the primary research questions has been so compromised by the restrictions being placed on the project that the results will have little scientific value. If the contribution of the study is compromised and conclusions limited, the investigator may decide that it is not productive to proceed with the research.

Securing Teachers' Cooperation for Research

The success of any and all research in school settings depends on the cooperation and support of teachers. Administrative support for a project, while important, often is not sufficient to gain the interest or cooperation of teachers, who ultimately have direct control over the researcher's access to students. For this reason, meeting with teachers separately from the school administration to explain the project to them in detail is critical to the successful implementation of research in school settings. This allows teachers to ask questions and learn about what time and involvement will be required from them to facilitate data collection.

To facilitate their understanding of the research, teachers should be given a brief summary of the research, which describes in concrete terms how and why previous research and/or clinical need had underscored the importance of the proposed study. Generating a list of previous findings and their implications, for example, a "fact" sheet can also be a useful and straightforward strategy to summarize such information. Giving teachers copies of all materials to be used is also important. In our experience, more experienced teachers generally express interest

in proposed measures and ask questions about why they were selected and what they are expected to measure. Moreover, teachers may comment on specific items that they thought were interesting and others that they believe should be removed.

As a follow-up to distributing the information about a proposed study, facilitating the teachers' discussion of the proposed measures and listening to their comments, questions, and concerns are useful ways to let teachers know what to expect in the research and to help them feel that they are an important part of the research process. Moreover, helping teachers to experience a sense of control over their decision to participate (i.e., that they are not being "forced" to participate by their administration) generally facilitates their involvement. In fact, in our experience, if they are invited to participate in the context of a dialogue about the proposed research, teachers often become more enthusiastic and supportive of a project with their students and may even offer additional assistance to researchers. Moreover, taking sufficient time to develop successful working relationships with teachers and administrators may pave the way for their collaboration in future projects. In conducting such dialogues with teachers, researchers should anticipate that they may ask some difficult questions, for example, why was this procedure chosen? To promote an atmosphere of collaboration, investigators should answer these questions and concerns directly and nondefensively.

Enhancing Student and Parent Participation in Research in School Settings

Once the researcher has secured the cooperation and support of teachers, facilitating the students' and their parents' participation in research becomes the next challenge. Because participation rates can be relatively low in some school-based studies, researchers should take active steps to ensure maximum participation. While parents make the sole decision concerning younger children's participation, middle and especially high school students should be given written information about the project and asked for their informed consent. Meeting with groups of students either in the classroom or in an assembly format is one useful vehicle for explaining the purpose of the study and what will be required of students. Once students have agreed to participate, gaining consent from parents is also generally required.

Securing Parental Consent for Participation

While Human Subject Review Boards at some universities may approve a passive parental consent procedure (i.e., parents must return a consent form only if they do not want their child to participate), other universities will not accept this procedure. In such situations, the researcher is faced with the challenge of mailing or sending parental consent forms to each child's legal guardian for their approval. In this case, all forms must be returned for a child to participate. This procedure, which ultimately protects both the researcher and the school, may also decrease the number of students who can legally participate. One procedure to facilitate parental consent is to include a letter of support for the project from the school and school system administration. Giving parents information that the proposed project is, in fact, valid and supported by important individuals in the school may increase parental approval and cooperation in research (Sieber, 1992).

While many researchers who conduct research in school settings are most interested in obtaining data from children, depending on the research question, it is sometimes important to gather data from parents as well. This objective is often easier to meet in a hospital or clinic setting because parents usually bring their children to their appointments. In school settings, however, access to parents is often gained only through mailings or special interest group meetings (e.g., Parent-Teacher Association). However, the latter method of eliciting parental participation can pose sampling problems because such samples are generally skewed toward highly involved and active parents.

SPECIAL ISSUES IN CONDUCTING RESEARCH IN CHILD WELFARE SETTINGS

Investigators who conduct research with special populations such as abused and neglected children and/or children in foster care need to understand the special requirements of research in child welfare settings. Understanding the specific mission of child welfare is necessary to anticipate challenges in conducting research in these service settings. The child welfare system in the United States exists to protect the needs and rights of children, by law investigates all mandated reports of child abuse and neglect, and makes a determination of the likelihood that abuse occurred. In order to ensure children's physical safety, the child welfare system often brings children into custody and places them in foster care or treatment facilities.

Developing an Agreement to Conduct Research

One common issue that arises in conducting research in child welfare settings that is not encountered in research in hospital and school settings is the need to obtain formal, legal consent from the child welfare agency. This is necessary because in many cases, as noted above, the agency is the child's legal guardian. Most researchers are accustomed to thinking about obtaining informed consent from the parents and, at an age-appropriate level, from the child. The legal status of child welfare as the custodial agent that must give informed consent for the child adds another difficult step to the process of obtaining consent from research participants and raises some difficult questions. These include: Who should be approached to give consent for children's research participation? How does one arrange for consent with a large system who maintains legal guardianship of children? Practically speaking, this may mean that a formal contractual agreement between the child welfare agency and the researcher is needed before any research can take place.

Before embarking on drafting a contract, the prospective researcher should take every opportunity to become known to the child welfare system and to relevant staff. Relationship building is every bit as important in establishing an effective research-related collaboration with staff in child welfare settings as it is in other settings. Researchers can utilize several different strategies to establish their visibility and/or utility to the child welfare system staff in individual communities. In our experience, the researcher's ability to find ways to demonstrate his or

her expertise and interest in issues that are helpful to staff of the child welfare system will often lay the groundwork for collaborative research. Demonstration of such interest can include the following strategies: (1) seeing families for assessment, providing consultation, assessment treatment and referrals; (2) being available to provide expert testimony, which is helpful to workers and administration; and (3) joining debate in one's local community about child protection and family preservation.

Managing Research in Child Welfare Settings

Much has been written (US Advisory Board on Child Abuse and Neglect, 1991) about the overwhelming demands on time and energy of social workers who provide services to children and providers in this system. Consequently, as is true in conducting research in other clinical settings, it is imperative for researchers to secure sufficient resources to allow them to collect their data independently without adding to an already strenuous workload of the child welfare staff.

It is also helpful for researchers to find ways to support developing partnerships with agencies through formal recognition of their collaborative roles on a project and developing resources to support these roles. Experts in child welfare research (Hoagwood, Landsverk, Nelson, & Horwitz, 1994) recommend structuring investigators' roles on a research project to enable a collaborating child welfare agency to designate one of their leaders as a co-investigator. In addition, ensuring that one's research budget covers any costs that may be incurred by the child welfare agency in working on the research is another helpful way to facilitate the implementation of research in such settings.

As is true of research in any clinical setting, various practical problems that limit data collection in child welfare settings may arise after the project has been designed and relevant legal obligations have been met. For example, an investigator initially may have proposed that a research assistant would go to the child welfare agency to read files and code relevant variables. However, the investigator may subsequently find that in some settings researchers who are not employees of the child welfare agency are denied legal access to the files. To solve this problem, one could either decide to gather the information directly as part of an interview process with clients of the child welfare system or try to arrange for someone who has the legal access to serve as research assistant for that portion of the project that involves reading files.

As this example suggests, working with large bureaucracies such as child welfare agencies takes time and can be frustrating. Nevertheless, with a thorough working understanding of the mission of the child welfare system and its day-to-day operation, we have found that solutions to many of these problems are possible. One example of effective management of a difficult problem in data collection in research with the child welfare system was a study of the outcomes of a cohort of children who were seen for evaluation and treatment by the child welfare agency for alleged sexual and physical abuse. The most significant challenge in conducting this study was to locate the children, many of whom were formerly in the custody of the child welfare agency. An extraordinary number of attempts were made, which eventually proved to be successful, to contact caseworkers of children who were still in the care of the child welfare agency to ask their help in locating families.

CONDUCTING RESEARCH IN THE CHILD MENTAL HEALTH SYSTEM

Clinical child psychologists who are interested in research with clinical populations, including research concerning program evaluation, will sometimes conduct research in the mental health system. The mission of the child mental health system is to promote the effective functioning of children and their families at home, at school, with peers, and in the community. The child mental health system is under increasing pressure from managed care companies to demonstrate its effectiveness. Consequently, there is a compelling need to document the effectiveness of clinical services that are actually delivered in mental health agencies and systems that work with children and families in community settings (Bickman et al., 1995; Kazdin, 1996; Weisz, Donenberg, Han, & Weiss, 1995).

Developing Collaborations with Mental Health Agencies for Children

There are several obstacles to developing research-related collaborations to mental health agencies for children that investigators should consider. For example, researchers who work with staff at community-based mental health agencies need to understand economic factors that might cause their potential collaborators to be distrustful of research. Agencies who provide services to children in families in community settings are often dependent on various sources of funding, such as state and local government and foundations, to support the services that are provided in these settings. Funders are often interested in supporting those services that have been found to be effective. On the other hand, providers may be motivated to continue services that they believe are effective based on their clinical experience, even though these services may not have been shown to be effective based on research. Consequently, mental health agencies who agree to participate in research may be put under greater scrutiny by their funding sources compared with agencies that decline participation.

Given potential disincentives to conduct research for mental health agency staff, such as that described above, researchers who are interested in working with colleagues in such settings need to engage in extensive efforts to develop their trust and cooperation. Some examples of strategies to develop trust and collaboration that are especially relevant to mental health settings include: (1) bringing a new, empirically supported intervention to the mental health community to open the door to conduct research with the mental health board; (2) providing consultation under contract to a department of mental health on issues of primary interest concerning service delivery for children; (3) helping the agency to redesign mental health services for abused children after funding limitations changed available methods of service delivery; and (4) giving in-service talks to a consortium of mental health board representatives on issues relevant to their agenda.

Forming partnerships with mental health agencies requires a great deal of sustained effort and evolves slowly over time. We have found that some of the best ways to demonstrate one's utility to a mental health agency are to follow the agenda and research questions that are of primary interest to the group whose population you hope to involve in research, to be willing to work with staff in the system on issues of concern to them that may be unrelated to the research, and to be both knowledgeable and respectful of the constraints under which the system works.

Barriers to Conducting Research in Mental Health Settings for Children

What are the most likely barriers in conducting research in mental health systems for children are encountered and how can they be resolved? We have identified the following specific barriers in conducting such research: (1) securing approval for the project from necessary collaborators; (2) problems that limit implementation of research as is proposed; and (3) managing the impact of changes in services that are being studied. In the next several sections, we describe examples of these problems and strategies to solve them.

Difficulty Securing Approval and Endorsement of the Project

It may be very difficult to obtain necessary approval and endorsement for some projects. For example, we encountered problems in obtaining approval for a study that was designed to predict which allegedly sexually abused children would develop sexually offending behavior as a consequence of their abuse. To locate these children, we needed the help of a number of agencies, including mental health agencies. Despite the endorsement of the project by the agency, it took literally a year to gain the needed legal approval. A concerted effort, which involved extraordinary patience, persistence, and time in working with the agency staff on other projects, eventually resulted in a signed contract and the ability to contact guardians of children for the research.

Problems That Limit Implementation of Research as It Is Proposed

Problems also can arise in implementing projects in mental health settings in accord with a researcher's proposed design. For example, a study that compared an innovative mental health intervention with a "usual services" condition within a mental health agency raised significant staff concerns about feasibility of assigning of cases and the potential of demonstrating that the services that were currently being delivered were inferior to the new model. These concerns were resolved by employing an additive design in which the customary patterns of service were compared with an intervention model that combined the customary pattern of services with the best practices intervention.

Ultimately, despite the above compromise, the above described research could not be implemented successfully. The mental health agency defaulted on its initial commitment to conduct the research. After working for a year to understand barriers and resolve them, the project could not continue. Based on the "lesson" that research with this particular agency did not prove to be feasible, a project now including random assignment to treatment groups was implemented in a neighboring county under the auspices of the mental health board.

Managing the Impact of Changes in Services That Are Studied

Researchers who collaborate with agency staff concerning studies of service delivery do not always have control over how services that are provided and/or structured. Unexpected changes in patterns of services that are being studied pose very difficult methodological problems for research that is conducted in mental health settings with children. The problem of contamination of control interven-

tions with experimental treatments is a potential issue in any intervention research project but is especially common in research conducted in complex settings such as mental health agencies. For example, Burns, Farmer, Angold, Costello, and Behar (1996) have described how a usual services condition can easily evolve into a more intensive interventions such as if control case managers, working alongside the "enhanced" case managers, adopt their techniques.

Strategies for Conducting Research in Mental Health Settings with Children

Based on our experience, strategies that may prove to be of practical utility when conducting research in mental health settings (or other complex systems with children) include: (1) make informed decisions about continuing research in the system; (2) recognize one's strengths and limitations; and (3) develop advocacy for one's research, both inside and outside the system.

Make Informed Decisions about Continuing Research in the Setting

Researchers need to recognize that even projects that are blessed by good timing, staff who respect each other, and adequate funding may encounter obstacles that are truly insurmountable. At such times, it often is best to modify the primary focus of a project and try another approach or even another setting. The researcher's ability to develop guidelines about when it is best to persist in implementing research as designed versus when to modify or even abandon a project is an important attribute.

Recognize One's Strengths and Limitations

Each researcher has individual strengths and limitations as well as a threshold level of tolerance for managing the various problems that can arise in conducting research in service systems for children. To be effective in conducting research in systems that provide services for children and families, researchers need to appreciate their own blend of patience, resources, and limitations. For example, one researcher may not mind having to collect data her- or himself when assistants are not available (such as during the holiday season), while another may not be able to make this type of commitment. One researcher may have special skills in writing proposals but may become anxious in meetings with heads of systems to negotiate roles on the project. Anticipating research tasks that are likely to create difficulties and making a plan about how to manage them is a wise strategy. Given the complex nature of research in service settings, teams of researchers often are needed to divide and share research tasks in ways that best take advantage of individuals' special strengths.

Develop a Strategy to Advocate for One's Research Inside and Outside the System

One strategy that can be used to address the difficult obstacles encountered in child mental health system research is to form partnerships with individuals who have special expertise that can inform the project. These individuals may work in

other parts of the system and/or in other settings in the community. Such internal–external systems partnerships (Timmons-Mitchell & Kloker-Young, 1993) can help to provide both the perspective and energy to manage the difficult problems that are presented by research with children in child mental health agencies. For example, in implementing a clinical intervention demonstration project in a child mental health agency, the issue of supplying beepers and cellular phones created a major stumbling block to the research. Because the intervention was designed to provide an intensive, around-the-clock response to a mental health crisis, providing therapists with the ability to be accessible to their clients was critical to implementing the intervention. In this situation, the investigator appealed to the experiences of other influential professionals outside the setting who had successfully conducted a similar intervention program in other settings. Using this information to show how therapist accessibility had been critical to the success of the previous intervention ultimately convinced agency administrative staff to supply the beepers and phones.

FUTURE DIRECTIONS TO ENHANCE RESEARCH WITH CHILDREN IN CLINICAL AND APPLIED SETTINGS

In the course of preparing this chapter, we were impressed with the formidable challenges faced by researchers in applied settings with children to ensure that their research is implemented in ways that do not compromise methodological integrity or ethical standards. For this reason, we believe that information about the special problems involved in conducting such research and strategies to surmount them should be disseminated and used to train researchers in clinical child and pediatric psychology. Several suggestions to accomplish these aims are now considered.

Disseminate Information Concerning Problems and Strategies of Conducting Research in Applied and Clinical Settings That Serve Children and Families

Researchers need to disseminate information concerning the special problems and strategies involved in conducting research in a range of applied and clinical settings that serve children. For this reason, it is important that they keep detailed records of the problems that arise during the course of data collection and decisions that were made to manage them. Such information not only is important to help researchers to interpret published findings but to those who may want to replicate the procedures.

Because information concerning the pragmatic problems in conducting research in settings with children is rarely included in published accounts, researchers and graduate training programs may want to develop manuals to guide data collection based on their experiences in specific settings. For example, program directors of training programs could request that students document the successful strategies they used to collect their data, as well as the hurdles and barriers that they faced in a department-wide compendium. Such program-specific handbooks also could include important anecdotal information on the experiences in successful and unsuccessful research projects, as well as specific materials that

would help students initiate studies (e.g., copies of IRB materials from local clinical settings, hospitals, schools).

Train Students to Conduct Research in Clinical Settings with Children

Our experiences also underscore the need for specialized training methods to help students to conduct research in applied and clinical settings that serve children and families. Potential training experiences include training in conducting pilot work, thesis and dissertation research, and formal course work in applied research methods.

Train Students to Conduct Pilot Work to Refine Data Collection Procedures

Unless they have collected data in clinical settings with children, investigators often cannot anticipate the specific problems that can occur and the necessity to change procedures to improve the quality of their data collection. For this reason, training students to conduct pilot work to develop and modify data collection procedures in well-planned pilot studies is critical to anticipate potential problems and to modify procedures to facilitate the efficiency and validity of data collection. It is much more effective to utilize a planned pilot study to develop and refine a workable procedure for data collection that can be applied throughout the study rather than it is to make substantial modifications in procedures after the project has begun.

Conduct Thesis and Dissertation Research in Clinical Settings

Training students to conduct a thesis or dissertation research in hospital, school, mental health, or child welfare settings provides important learning opportunities. Student researchers with a strong personal commitment to collect data in applied clinical settings should work with mentors who are well known to relevant staff and they should develop strong working relationships with key individuals in these settings. If possible, establishing working relationships through clinical placement, externship, or volunteer work can help to promote a student researcher's visibility in clinical settings. Moreover, working with someone who is already conducting research in such settings and/or is employed by the setting is strongly recommended. For example, developing mentorship teams, which include a faculty member in the department of psychology and a senior staff member at an applied service agency, can facilitate the conduct of masters and dissertation research in clinical service settings. This model proved to be effective in facilitating a dissertation project that tested the efficacy of videotaped preparation for parents on attendance and attitudes concerning their children's psychotherapy (Ludwig, 1997). One of us (Drotar) provided faculty supervision for the development of the project, while another clinical faculty member (Jeremy Shapiro), an experienced researcher and senior staff member at the mental health service agency, provided the supervision and institutional support that was in many ways most critical for the success of this student's research. Students clearly benefit from opportunities to discuss the issues and problems in collecting data in didactic and applied clinical settings with their mentors.

Develop Formal Course Work in Applied Research Methods

Formal course work and didactic training can also be utilized to help students appreciate the pragmatic problems involved in conducting research in clinical settings, the common threats to validity encountered in such research, and ways of managing them. For example, a semester-long course on applied clinical research methods includes lectures that summarize and describe the common threats to the validity of research in applied settings, problems that are encountered in data collection, and ways of managing them in study design and data analysis. We have involved researchers who have conducted successful research in a range of applied settings with various clinical populations, for example, very-low-birth-weight infants, children with conduct disorders, children with various chronic health conditions, and so forth, in this seminar. In addition to presenting information concerning the design and findings of their research, these researchers were asked to describe how they managed common logistical problems related to data collection, subject recruitment, and attrition, and so on, in their research. Students have gained much from opportunities to discuss these issues with experienced faculty researchers.

Enhance Continuing Education Opportunities in Applied Research Methods

Continuing education courses, in which experienced researchers describe the problems in conducting research in applied clinical settings with children and the strategies of managing them would also be very useful, not only for students but for experienced researchers. Ideally, these would take the form of interactive roundtable or symposium formats. Finally, working consensus conferences in which experienced researchers develop recommendations for strategies to address the salient logistical problems in conducting research in clinical settings with children and families would be not only instructive but would help to move research forward.

REFERENCES

- Bickman, L., Guthrie, P. R., Foster, E. M., Lambert, E. W., Summerfelt, W. T., Breda, C. S., & Heflinger, C. A. (1995). *Evaluating managed mental health services: The Fort Bragg experiment*. New York: Plenum Press.
- Bond, L. A., & Compas, B. E. (Eds.). (1989). *Primary prevention and promotion in schools*. Newbury Park, CA: Sage.
- Brody, G. H., & Sigel, I. E. (Eds.). (1990). *Methods of family research. Biographies of research projects. Vol. II, Clinical populations*. Hillsdale, NJ: Lawrence Erlbaum.
- Bull, B. A., & Drotar, D. (1991). Coping with cancer in remission: Stressors and strategies reported by children and adolescents. *Journal of Pediatric Psychology*, 16, 767–782.
- Burns, B. J., Farmer, E. M. Z., Angold, A., Costello, E. J. & Behar, L. (1996). A randomized trial of case management for youths with serious emotional disturbance. *Journal of Clinical Child Psychology*, 25, 176–186.
- Cicchetti, D., & Manly, J. T. (1990). A personal perspective on conducting research with maltreating families: Problems and Solutions. In G. H. Brody & I. E. Sigel (Eds.), *Methods of family research. Biographies of research projects. Vol. II, Clinical populations* (pp. 87–134). Hillsdale, NJ: Lawrence Erlbaum.
- Drotar, D. (1989). Psychological research in pediatric settings: Lessons from the field. *Journal of Pediatric Psychology*, 14, 63–74.

- Drotar, D. (1993). Influences on collaborative activities among psychologists and physicians: Implications for practice, research and training. *Journal of Pediatric Psychology*, 18, 159–172.
- Drotar, D. (1995). *Consulting with pediatricians: Psychological perspectives*. New York: Plenum Press.
- Drotar, D., & Robinson, J. (1999). Researching failure to thrive: Progress, problems and recommendations. In D. Kessler & P. Dawson (Eds.), *Failure to thrive in infants and children: A transdisciplinary approach to nutritional adequacy in children* (pp. 97–98). Baltimore: Paul H. Brookes.
- Hoagwood, K., Landsverk, J., Nelson, C., & Horwitz, S. (1994, March). *Mental health services from the child welfare perspective: Recent findings and methodological challenge*. Paper presented at the Seventh Annual Research Conference A System of Care for Children's Mental Health: Expanding the Research Base. Tampa, FL: Research and Training Center for Children's Mental Health, Florida Mental Health Institute, University of South Florida.
- Kazdin, A. E. (1996). Combined and multimodal treatments in child and adolescent psychotherapy: Issues, challenges and research directions. *Clinical Psychology: Science and Practice*, 3, 69–100.
- Ludwig, A. (1997). *Impact of videotaped preparation on the attendance of children and families in psychotherapy*. Unpublished doctoral dissertation, Case Western Reserve University, Cleveland, OH.
- Meyer, A., Miller, S., & Herman, M. (1993). Balancing the priorities of evaluations with the priorities of the setting: A focus on positive youth development programs in school settings. *Journal of Primary Prevention*, 14, 95–113.
- Seiber, J. E. (1992). Planning ethically responsible research. In *A guide for students and internal review boards*. Newbury Park, CA: Sage.
- Timmons-Mitchell, J., & Kloker-Young, C. (1993). Three coins in a fountain make good sense for children. In H. K. Cleminshaw (Ed.), *Home-based intervention: An interdisciplinary approach to programming and research* (pp. 157–166). Akron, OH: The University of Akron. Ohio Department of Mental Health and the Center for Family Studies.
- US Advisory Board on Child Abuse and Neglect. (1991). *Creating caring communities: Blueprints for an effective federal policy on child abuse and neglect*. Washington, DC: US Government Printing Office.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995). Bridging the gap between lab and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63, 688–701.

13

Multi-Institutional and Multidisciplinary Research Collaboration

Strategies and Lessons from Cooperative Trials

F. DANIEL ARMSTRONG and DENNIS DROTAR

Over the 30 years since the formal establishment of the field of pediatric psychology, a number of transformations have occurred that have affected the nature and focus of research in this field, some of which also are applicable to the field of clinical child psychology. Early publications in the *Journal of Pediatric Psychology* were based on observations described in case reports or single-subject designs. These reflected the growing awareness of the potential unexplored relationships between the behavioral and medical aspects of various diseases and the procedures used to treat them. As the field matured, research drifted away from the focus on individual cases and began to rely more on group designs and methodology. This focus resulted in a new set of difficulties, characterized by small, potentially nonrepresentative samples, broadband research with little consideration of specific disease, treatment, or developmental variations and restrictions on the types of research possible due to limited access to populations. In an effort to solve these problems, current efforts in pediatrics, including pediatric psychology, are moving toward increased research collaboration, both across institutions and across disciplines, although at a somewhat slow pace. Of the 186 empirical articles published in the *Journal of Pediatric Psychology* between January 1993 and August 1997, only 10 (5%) involved collaborations that included data collection at multiple institutions. Thus, while there is evidence of multisite collaboration occurring in the

F. DANIEL ARMSTRONG • Department of Pediatrics, University of Miami, School of Medicine, Miami, Florida 33101. **DENNIS DROTAR** • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

field, it is still in its relative infancy. To promote a better understanding of this approach, it is our intention to identify several models of multisite collaborative research, pointing out the benefits and pitfalls that may be encountered.

There is a critical need for an increased focus on multisite collaborative research and several important reasons for this focus. The first and most important is the need to adequately address the specific nature of problems and outcomes for children with relatively low-incidence medical problems. For example, children treated for Wilms' tumor (cancer of the kidney) receive a significantly different course of therapy, have a different prognosis, and experience very different late effects of treatment than a child treated for acute lymphocytic leukemia with central nervous system disease. Consequently, the types of assessment and surveillance needed, as well as the type and intensity of behavioral intervention needed, are likely to be substantially different for these two groups of children. Unfortunately, single-site studies typically have looked at children with cancer, ignoring the disease and treatment variations because there were not enough subjects available to permit disease-specific analysis (Armstrong, 1995). The difficulties resulting from this approach are twofold: First, global studies may fail to identify problems specific to subpopulations of children, leading to lack of specific interventions for those children who indeed experience various complications. Second, global studies may identify problems for children with cancer that do not apply to children with specific types of cancer, leading to their receiving unneeded services and overutilizing limited resources. Cooperative, multisite research endeavors represent the only mechanism for conducting research on low-incidence diseases that permits the level of analysis that can prevent each of the above problems from occurring.

Despite the obvious need for collaborative research, most psychologists are poorly trained in the specific skills and knowledge necessary to conduct these studies. Multidisciplinary and multicenter research requires the kinds of political, logistic, communication skills and attention to detail that do not exist in any single investigator study. There also are a number of barriers to multisite collaborative research that apply to all investigators, not just psychologists, but these may not be recognized until they bring a project to a grinding halt (Smyth et al., 1994).

In this chapter, we will identify training issues, skill issues, and anticipated barriers along with possible solutions for individuals who conduct collaborative research. The examples used to illustrate the points of this chapter are based on our experiences with the limited institution collaborations with colleagues at two to three other institutions, as well as with the Pediatric Oncology Group, Cooperative Study of Sickle Cell Disease, Pediatric AIDS (acquired immunodeficiency syndrome) Clinical Trials Group, the Diabetes Control and Complications Trial (DCCT), and other multisite studies. We will also use these experiences to point out pitfalls, demonstrate possible solutions, and make recommendations for future directions.

MODELS OF COLLABORATIVE RESEARCH

There are multiple ways that research can be conducted using a collaborative model. Within institutions, investigators who share a common interest and back-

ground may develop collaborative research laboratories. Historically, this has been restricted to intradisciplinary efforts where, for example, two psychologists or two pediatricians work together, usually at the same institution or laboratory. In recent years, technology has made communication between investigators at different institutions more efficient and intradisciplinary collaboration across multiple institutions has become more common (e.g., Drotar, Agle, Eckl, & Thompson, 1996; Quittner et al., 1996; Sargent et al., 1995). As the complexity of interactions between biological, social, and behavioral factors has been increasingly recognized, collaborations between investigators of dissimilar backgrounds have begun to develop. This has led to a number of interdisciplinary collaborations between pediatricians and pediatric psychologists, illustrated by research in childhood cancer (e.g., Kazak et al., 1996), sickle cell disease (Lemanek, Horwitz, & Ohene-Frempong, 1994), and human immunodeficiency virus (HIV)/AIDS (Hardy, Armstrong, Routh, Albrecht, & Davis, 1994). In the following sections, we will review models of collaboration between investigators of different disciplines, notably psychology and pediatrics, who are located at the same institution. We will then examine models using collaboration between investigators of various disciplines who work across institutions within the framework of limited institution agreements or within the context of large, multicenter investigational trials. Each model offers both opportunities and potential pitfalls, which are identified.

MULTIDISCIPLINARY COLLABORATION WITHIN INSTITUTIONS

Models of Collaboration

Within institutions, multidisciplinary research collaboration between pediatric psychologists and pediatricians (and subspecialists) typically follows one of three models. In the first model, serving as the principal investigator (PI), the psychologist poses a primary behavioral question that relates to pediatric care and elicits the involvement of a pediatrician as a contributor to the project. For example, a graduate student in pediatric psychology decided to investigate whether there were differences in behavior, preference, and physiological arousal in children receiving immunization injections in a well-child clinic. A team of investigators consisting of pediatric psychologists, a health psychologist with experience in psychophysiological measurement, a primary care pediatrician, and a pediatric nurse systematically investigated this question, leading to the observation that children's distress increases when their mother is present for the injection, but their preference is for the mother to be in the room (Gonzalez, et al., 1989). The results of this study led to changes in the clinical practice in the clinic. In this type of collaboration, the research question is behavioral, and the pediatrician is asked to provide technical or measurement assistance.

A second research model reverses the roles, with the pediatrician posing the primary question and serving as the PI for the project and the psychologist providing technical support as a contributor to the primary question of the study. Several years ago, our pediatric hematologists became interested in the actual impact of prophylactic penicillin on bacterial infections in our children's hospital. One of the important variables was whether the children were taking the penicillin, so medi-

cation adherence was a concern. The resulting study included a psychologist for the purpose of assessing compliance in what was otherwise a medication treatment study (Pegelow, Armstrong, Light, Toledano, & Davis, 1991). In such cases, the research question is biological, and the psychologist is asked to provide technical or measurement assistance only.

A third research model exists when the behavioral and biological questions are integrally related and of relatively equal importance. In this case, the psychologist and pediatrician serve as co-investigators, emphasizing shared responsibilities for all aspects of the research. This model requires shared knowledge and concepts between the investigators, with the pediatrician having a clear perspective on the behavioral aspects of the question and the psychologist having a similar understanding of the biological aspects of the study. In such research, the line between what is behavioral and what is biological can become blurred as both investigators integrate their knowledge in pursuit of a question of importance to both. For instance, the effects of HIV on young children's development is a critical concern, both because of the developmental outcome and need for special services that the child may have and because changes in development may be markers for significant changes in the course of the disease.

Benefits of Local Collaboration

Irrespective of the collaborative model used for a particular research project, there are a number of benefits of psychologist–pediatrician collaboration. Proximity increases ease of collegial contact, improving the opportunity for the investigators to interact on a regular basis and maintain project momentum. Such contact also may pay dividends as professionals from different professional backgrounds become better educated about the skills and knowledge base that each bring to the research endeavor. In our experience, the benefit for psychologists involved in collaboration with pediatricians in medical settings is substantial. In many cases, psychology is initially viewed as a soft science with little to offer in a pragmatic way to research on physical illness. Close collaboration often (in fact usually) debunks this myth, increasing the value placed on both psychological methodology and on the behavioral issues associated with disease, and may even aid in the translation of basic science applications into clinical practice (Morrow & Bellig, 1994).

The working relationships that originate from collaborative efforts often generalize beyond the specific research project to clinical practice and education. For example, referrals, requests for psychological consultation in primary and acute care settings, and integration of psychologists into multidisciplinary clinical teams frequently are the result of research collaborations (Drotar, 1995). Once this occurs, collaborative efforts in education often develop, with psychologists being asked to actively participate in medical student, resident, and fellow education in both clinical and research arenas. All these activities serve to enhance the role of psychology in the medical setting, increase the value of the research produced by psychologists, and facilitate the use of behavioral findings in treatment planning and policy development. An unexpected benefit is the opportunity for psychologists to learn more about the pathophysiology and treatment of the illnesses that give rise to the need for behavioral research, to develop assessment and treatment models that are more integrative with medical information, and to have pediatri-

cians take an active interest in assisting in the training of psychology graduate students, interns, and fellows (Drotar, 1995).

At a practical level, many other potential problems can be avoided or alleviated through a collaborative approach to research. Institutional review board approvals, development of comprehensive approaches to informed consent, access to populations and easier subject recruitment, and standardization of study procedures, data management, and manuscript preparation are simplified by proximity, trust among co-investigators, and shared understanding of the research question, its relevance to clinical practice, and its ease of implementation (Crooks, Colman, & Campbell, 1996; Lantz, Reed, & Lewkowitz, 1994).

Limitations of Local Research Collaboration

There are, however, several limitations to the institutional collaboration model that may be difficult to overcome. For conditions that are relatively rare, most institutions are limited in the number of patients available for study. In these cases, research questions also may be limited in several ways. First, the total number of subjects available may provide insufficient power to detect statistical significance, particularly when effect sizes are small (Brocklehurst, Elbourne, Garcia, & McCandlish, 1996). Second, while there may be an overall adequate sample size, the sample may be so heterogeneous that the research questions can only be very general in nature (Armstrong, 1995). Evaluation of significant covariates, interactions, developmental factors, and nested treatment effects thus becomes quite difficult. Third, limited samples may lead to the population being overstudied. In such situations, children with rare conditions become "professional research subjects," recruited for every medical and behavioral study that is conducted at that institution. This strategy runs the risk of developing an extensive knowledge base with a sample that is potentially nonrepresentative of children with a particular disease who are treated at other sites, as well as promoting a response bias in these subjects that results from repeated participation (Drotar, 1995).

A second limitation of the institutional collaboration model may be difficulties in obtaining either the technical expertise or equipment necessary to answer the question of interest. Whether the need is for an expensive piece of integrated computer equipment, neuroimaging technology (e.g., positron emission tomography), or measurement skills that require specialized training, costs may preclude any single institution from being able to provide the infrastructure necessary for the study. This sometimes leads to research that is conducted to meet an investigator's specific need (e.g., completion of a dissertation, publication for tenure) rather than to further the state of the art in methodology and measurement, and hence promoting advances in the field.

MULTI-INSTITUTIONAL COLLABORATIVE RESEARCH

When the resources or collaborative relationships for conducting meaningful studies do not exist at a single institution, multisite and multicollaborator research represent a viable and exciting alternative. Two models of multi-site collaboration have emerged, neither of which is uniquely superior to the other. In fact, the two

models—limited institution collaboration and large multicenter cooperative group research—each involve different assets and limitations. However, these assets and limitations have led to complementary research approaches, where limited institution collaborations serve as pilot projects that are then instituted as a multi-center clinical trial within the context of a large cooperative group. Additionally, the large collaborative group may develop clinical trials with specific, highly technical questions that can be addressed by a smaller subset of institutional investigators using the limited institution model. Therefore, as these models are described, it is important to maintain a complementary or cooperative perspective rather than a competitive one when evaluating the strengths and weaknesses of these approaches.

Limited Institution Collaborations

As in the local institution setting, there are several models for investigators from more than one institution to collaborate on research on children's health conditions. In the limited institution collaboration, investigators at multiple institutions agree to share resources to conduct a study that could not be carried out at any of the individual institutions, usually because of inadequate subject samples or inadequate resources. Such studies may be inter- or intra-disciplinary, as in the local collaboration model. The basis for the collaboration may be similar populations across sites, with data collection occurring in a similar fashion at all participating institutions. This model has been used to develop measurement tools (Quittner et al., 1996; Snyder et al., 1997), assess the perceptions of cancer by siblings of children with cancer (Sargent et al., 1995), describe problem-solving strategies of children with diabetes (Thomas, Peterson, & Goldstein, 1997), and study psychological distress in children with hemophilia and their mothers (Drotar et al., 1996). Alternatively, it may involve sharing of site-specific resources, with data collection occurring at one or two sites, and other institutions providing special technological or data management support for the entire study, as in neurobehavioral follow-up of children with intraventricular hemorrhage whose mothers participated in a study of determinants of responsiveness (Onufrik, Saylor, Taylor, Eyberg, & Boyce, 1995). Other examples could include a study of differences in disease severity and the impact on psychological adjustment in children with different genetic variants of a disease. When sophisticated technology is necessary, such as functional neuro-imaging during cognitive tasks, subjects might be recruited at several participating institutions but undergo the actual study at the institution with the equipment and technical expertise. Each approach offers the opportunity to conduct research that could not be carried out by investigators at a single institution.

Large Multi-Site Cooperative Groups

Over the years, the National Institutes of Health (NIH), the National Institute of Mental Health (NIMH), and more recently several private foundations recognized that advances in scientific knowledge in low-incidence disorders cannot be obtained by individual investigators working in a single institution setting. Subject populations for many clinical trials simply did not exist at single institutions to support research or any topics other than broad questions, and even these types of

studies were open to criticism because of subject selection bias (Brocklehurst et al., 1996). Therefore, beginning in the late 1960s and early 1970s, the NIH began funding large collaborative groups, charged with the task of developing studies to address natural history and clinical trials questions that could not otherwise be conducted. In pediatrics, the early cooperative groups followed models that had been developed for adults with cancer, and eventually the Children's Cancer Group (CCG) and the Pediatric Oncology Group (POG) were formed to conduct clinical trials devoted to treatment of childhood cancer (Pediatric Oncology Group, 1992; Stowe et al., 1984). Each group involves 80–100 institutions that work together to carry out protocols developed by the group and monitored by oversight committees at the NIH.

Over the years, each of these groups have developed psychology discipline committees that provide data on acute and late effects of treatment, as well as develop assessment and intervention protocols related to maintaining and improving quality of life of children treated for cancer. Using this model, several other cooperative groups have been formed that focus on low-incidence diseases in children to better describe the natural history of a disease or developmental outcome [e.g., Hemophilia Growth and Development Study (Hillgartner et al., 1993; Loveland et al., 1994); Cooperative Study of Sickle Cell Disease (CSSCD) (Farber, Koshy, & Kinney, 1985; Gaston et al., 1987); and the Diabetes Control and Complications Trial (DCCT) Research Group (1993)]. Other groups have been formed that focus on conducting clinical trials for treatment [e.g., Pediatric AIDS Clinical Trials Group (PACTG) (Rosendorf et al., 1993)]. In each case, the efforts of the group have been multisite, multidisciplinary, and usually prospective and longitudinal in design. Pediatric psychologists have been involved in various ways in the administration and scientific activities of each of these groups.

BENEFITS OF MULTI-INSTITUTIONAL COLLABORATIVE RESEARCH

Both limited institution and large multicenter group collaborations offer benefits that often cannot be obtained at a single institution, particularly research concerning uncommon medical and behavioral problems of children. The nature of the particular benefit is highly dependent on the specific research question, the organization of the group, and the personal relationships that exist among investigators.

Improved Methods, Including Sample Size

The most common benefit associated with multi-institutional collaborative research is decreased methodological limitations because of inadequate or non-representative subject samples. By using these collaborative models, investigators can define the research question, develop a power calculation, establish exclusion and inclusion criteria, and recruit a sample that will permit appropriate statistical analysis of the data obtained. The question of inadequate sample size is thus unlikely to be raised as a criticism of the study outcome. A second less obvious but perhaps more important benefit occurs when children with uncommon disorders (and their families) are permitted to participate in studies of relevance to their

disease and its treatment, particularly when the data from such studies result in improved identification, management, or cure of the disorder. In the absence of the more adequate sampling afforded in the multi-institutional model of research, the problems confronted by children with low-incidence diseases and/or behavioral problems cannot be studied with any methodological rigor (see Chapter 4, this volume).

Another benefit associated with multi-institutional collaborative research is advances in the knowledge base concerning treatment and clinical management. One potential problem in assessment and treatment of behavioral problems associated with uncommon pediatric diseases is overgeneralization of the nature and severity of problems faced by these children, based on limited observations of a few children who are referred for severe difficulties. Published reports of these problems based on limited observations may then lead to well-intentioned but intrusive efforts to intervene, often using interventions with little or no demonstrated efficacy or appropriateness. On the other hand, inadequate sampling of children with uncommon diseases may lead to a failure to identify critical problems that exist for a subset of these children. Carefully designed multi-institutional studies, either limited institution or large multicenter group, decrease the potential for such over- or undergeneralization to occur.

The recent experience of the Cooperative Study of Sickle Cell Disease (CSSCD) illustrates this point. In the 1960s, the question of whether children with sickle cell disease were at risk for learning problems was first raised. Chordokoff and Whitten (1963) conducted a single institution study with a limited sample of children and concluded that the children with sickle cell disease were not at risk for cognitive impairment. Further research on this topic did not occur for some 20 years, when several single institution studies in the late 1980s again suggested that there was a risk of cognitive impairment with sickle cell disease (Fowler et al., 1988; Swift et al., 1989). However, the samples in these studies were once again small, not truly representative of the sickle cell disease population and potentially confounded by issues of disease type, poverty, education of parents, and other socio-demographic factors.

To address this and other questions, the CSSCD launched a 19-center natural history study of a cohort of children with sickle cell disease recruited at birth. Beginning at age 6, and occurring every 2 years after that for a total of four evaluations, these children received a battery of cognitive tests, magnetic resonance imaging (MRI), physical examinations, and tests of pulmonary and other organ functioning. The large sample obtained ($N = 187$) of children with hemoglobin SS disease permitted design and statistical control of disease and socio-demographic factors that had previously confounded research in this area. The multidisciplinary approach (pediatric psychology, hematology, and neuroradiology) permitted grouping by disease variables, leading to the finding that 17% of children with hemoglobin SS disease had MRI evidence of brain abnormality and infarct, and the performance on tests of cognitive and academic functioning was worse for children with these brain abnormalities than for children with normal MRIs (Armstrong et al., 1996). This large, multicenter approach led to findings that could not have been obtained at any single center in the United States, and resulted in an answer to the question of cognitive impairment that identified a subset of

children, but not all children, with sickle cell disease as being at risk. The multi-disciplinary collaboration also led to interactive formulation of biobehavioral models of risk, not only due to brain infarction, but also due to chronic anemia and poor pulmonary functioning (Brown, Armstrong, & Eckman, 1993).

Increased Opportunities for Funding

A second potential benefit associated with limited institution or large multi-center cooperative trials is the increased opportunity to obtain funding for research. Potentially important and well-designed studies of behavioral functioning of children with chronic illnesses may not be funded because of inadequate sample sizes. The multisite cooperative research mechanism permits adequate sampling and also allows more precise and scientifically sound research to be conducted. Such models may be especially effective because they enable investigators to address complicated biobehavioral questions that have previously been impossible to study.

Scientific Diversity

A third, sometimes unrecognized, asset of multi-institutional collaboration is found in the interprofessional relationships that are developed within the collaborative structure. Multi-institutional collaboration permits investigators of similar interests with different strengths to share ideas, refine instrumentation, develop creative designs, and work together to maintain a high quality of data collection and integrity. Several examples from our own work may be illustrative. In 1987, one of us (DA) recognized a need for a measure of quality of life in children treated for cancer. A project was started at the University of Miami to develop such an instrument. At the same time, psychologists and oncologists in the Pediatric Oncology Group formed a special committee to address quality of life measurement in clinical oncology trials. The Miami Pediatric Quality of Life Questionnaire (Armstrong et al., *in press*) was one measure that was evaluated and included in the POG quality of life battery. Investigators at several other institutions incorporated this measure into their institutional batteries and provided input on item appropriateness and usefulness. Other investigators provided input on the use of the instrument with multicultural populations. The resulting measure has been systematically refined based on this experience and is now routinely used as a measure in POG clinical trial protocols examining outcomes following bone marrow transplantation and treatment of brain tumors in children.

Other examples of such collaboration may involve having a psychologist with strong interest in memory measurement working with a psychologist with strong interest in pain management on a study of the effects of chronic pain on memory development in young children. Pediatricians with an interest in anesthesiology also may participate, examining relationships between different types of sedating medications used to treat pain in young children. Likewise, an expert in statistical modeling may be recruited to aid in the development of a design that will permit this question to be answered. The likelihood of such a diverse set of investigators

being housed in a single institution with an adequate population to address the study question is quite small.

Opportunity for Interdisciplinary Learning Concerning Clinical Research

Much of our training as scientists focuses on developing expertise in relatively limited areas. However, in the field of pediatric health, it is becoming more widely recognized that many of the clinical problems of children are best understood as complex interactions between biological processes, learning, and current environmental contingencies. Unfortunately, current professional training models are often limited in interdisciplinary diversity. For example, pediatricians often have limited understanding of behavioral principles, despite the large empirical literature demonstrating the role of these principles in both observable behavior and biophysiological processes. On the other hand, most psychologists have limited exposure to specific information about childhood medical disorders, the medications used treat them, and behavioral manifestations of the side effects of these medications (Armstrong, 1995). Psychologists also may be unaware of sophisticated measurement techniques, for instance, functional imaging, advanced cardiac monitoring, and genetic marking, that may be of significant relevance to some kinds of psychological research. Multisite collaborative research provides unique opportunities for learning about technologies that are used by other professions in research and clinical care.

Coordinated Statistical Support and Data Management

One of the most valuable, yet limited, resources that is needed for high-quality, funded, biobehavioral research is biostatistical support. Despite the training of psychologists in the statistical method and statistical approaches, few psychologists are capable of maintaining current, state-of-the-art knowledge of statistical design and analysis in many areas of biostatistics. For example, training and experience with such methods as survival analyses and non-parametric statistical approaches is less frequently included in the psychology curriculum. Cooperative research endeavors provide the opportunity to utilize the advanced skills of biostatisticians in an efficient and cost-effective manner and at the same time increase the possibility of obtaining funding for projects, because most federal grants require the inclusion of an identified biostatistician or biostatistical consultant before funding is provided.

Efficient data management, data quality control, and database development are skills requiring special expertise. The multi-institutional collaboration model permits the inclusion of a knowledgeable and competent biostatistician as a co-investigator. Such an individual may elevate the overall quality of the project by bringing organized computer, data entry, and data management facilities and procedures to the study. Another important side benefit of such collaboration is the increased awareness and knowledge of the diversity of data management and statistical approaches available for analyzing complex data sets, particularly those involving longitudinal cohorts with variable attrition due to death or changes in disease status.

Benefits to the Patient

There are two direct benefits that seem to accrue to patients resulting from multi-institutional and multidisciplinary collaborative research. The first is a benefit to both the patient and to the investigators. Multi-institutional collaboration permits research questions to be answered much more quickly because of the increase in the sample size. Single institution studies of a phenomenon may require years to accrue adequate samples to address a given clinical question. Multi-institutional collaborative research may permit the same question to be answered in a fraction of the time. The benefit for the patient is more rapid availability of information relative to his or her care. The benefit for the investigator is more rapid completion of the study permitting future research to develop more rapidly.

A second benefit that accrues to patients relates to the speed with which research findings are disseminated and put into practice. Investigators working across institutions are likely to place contingencies on one another to publish findings. Moreover, the credibility of such collaborative work is likely to be high within the scientific and clinical practice community. This process eventually can facilitate wider distribution of clinically effective treatment to patients in need.

A third benefit to patients concerns the use of professional expertise to enhance patients' understanding of and adherence to research in the context of clinical trials. For example, it was essential to the success of the Diabetes Control and Complications Trial (DCCT research group) to enroll individuals with diabetes who were well-informed about and understood the trial's purposes and procedures and were relatively free of difficulties that would interfere with their adherence to clinical trial of the efficacy of intensive treatment of diabetes on the development and progression of long-term complications in diabetes. Mental health professionals, including psychologists and social workers, who worked in each site of the DCCT screened participants for a history of psychiatric problems, serious objections to the study by family members, or indications that a potential subject was following his or her treatment regimen poorly. Potential participants also were asked to complete several weeks of blood testing and record keeping as well as diaries of meals, exercise, and insulin regimens. In addition, patients were interviewed concerning their perceptions of the advantages and disadvantages of the trial and their willingness to be randomly assigned to treatments. These procedures helped to ensure that patients who were enrolled in the trial, including a subgroup of adolescents, sustained their participation over time.

Over the course of the trial, behavioral science investigators also worked with patients and families to evaluate the nature of adherence problems and provide interventions to promote adherence. The need for this type of intervention was anticipated by having at least one mental health professional at each participating site. Individual and family counseling provided by these professionals helped patients deal with a wide range of personal and environmental obstacles to adherence. Participants with difficulties with compliance were assisted by cognitive behavioral interventions such as time management, relaxation training, and behavioral contracts. The behavioral science professionals at each center also have facilitated staff interactions by their participation in team meetings and conferences. Finally, the behavioral science professionals met at least once a year over the

course of the DCCT to review their experiences and plan additional ancillary studies that enriched the overall scientific contribution of the research.

DIFFICULTIES OF MULTI-INSTITUTIONAL COLLABORATIVE RESEARCH

Despite the many clear-cut benefits to conducting research within the multi-institutional and multidisciplinary model, some difficulties are encountered that should be anticipated. These include limitations related to the scientific aspects of the study, problems posed by individuals with competing agendas and professional issues related to study responsibility and authorship. In the following section, we will describe some of these difficulties and possible solutions.

Scientific Limitations

While multi-institutional collaborations offer a number of benefits through access to larger sample sizes and diversity of knowledge in skills, they may be limited in the specificity of research questions that can be addressed, comprehensiveness of measurement and quality control.

Variations in Availability of Measures

Our experience in large cooperative trials has been that many institutions lack all but the most basic cognitive assessment devices, and measurement tools for infants and young children may be completely unavailable. Even when these tools are available, the question of adequate reimbursement for professional time, either through grant support of time or fee-for-service reimbursement, is a substantial obstacle to data collection at many sites. Studies utilizing certain sophisticated measurement approaches (e.g., computerized testing, psychophysiological assessments, or behavioral observation systems) may be impossible to conduct in the large, multicenter trial model. This problem is twofold: Without adequate funding, many centers are unable to provide the equipment needed to conduct the study measurements. Even if funding for equipment is available, appropriate training in the collection of study data is necessary but difficult to arrange (Crooks et al., 1996; Smyth et al., 1994).

Difficulty of Maintaining Standardized Data Collection

Even in a single institution study, investigators must be very concerned about issues of staff training and ongoing calibration of measurement, and this issue is magnified in the multiple institution research setting (see Chapter 12, this volume). Consequently, special procedures must be developed to maintain standardization of measurement across multiple institutions. Aggravating this problem is the fact that, even in the multi-institutional trial where more rapid acquisition of subject samples is possible, data collection occurs over time. Thus, there is a need for ongoing recalibration to prevent systematic drift or inaccurate measurement across time and settings.

Issues related to staff attrition that may add to the inherent difficulties of maintaining standardized measurement also must be addressed in multi-institutional research. Staff turnover within a single institution may cause short-term delays that can be addressed through institutional hiring and training procedures. In the multi-institutional setting, staff changes occur frequently and at different times across the institutions. Training staff at the beginning of the study may help to address this issue, but procedures for maintaining ongoing training are critical to the outcome of the multi-institutional trial.

Maintaining Quality Control

There are multiple opportunities for errors in data collection, transformation, and scoring in multisite collaborative research. For instance, in the early phases of the CSSCD investigation, the question of the accuracy of data in the CSSCD database was raised prior to preliminary statistical analyses. At this point, raw data collected at all participating centers were forwarded to a single institution where quality reviews were conducted. These consisted of rescoreing all data submitted and comparing these data to those originally submitted by each of the participating institutions and to the data that had been entered into the database. Overall, nearly 40% of the data in the CSSCD database were inaccurate. The errors originated at multiple points in the data collection process. Some mistakes were relatively minor and involved problems with age calculations of the children participating in the study. Others involved basic calculation errors (e.g., adding up incorrect raw scores) or standard score calculation mistakes. Surprisingly, in a number of cases, the wrong normative data were used, and in some cases scoring was accurate but transcription of data onto summary forms was incorrect. Perhaps because these data were managed by a professional statistical consulting center, the number of errors entered into the computer database were minimal. This observation led the CSSCD to initiate ongoing quality assurance review of all data submitted for this study. With this review and feedback to participating centers, the frequency of errors in data submitted dropped to below 3%.

Strategies to Reduce Scientific Limitations

Several strategies may be useful in reducing the methodological problems that are associated with data collection at multiple sites. First, proposals for funded research should include as line items purchase of identical equipment for all participating institutions. In the case of a large, multicenter study, these costs may be prohibitive. An alternative strategy is to purchase several sets of equipment maintained by a central study coordinator. As subjects are enrolled on study at different institutions, the institutional coordinator can notify the central study coordinator of the registration, and measurement equipment can be shipped (with return shipping provided) to the institution as needed. This strategy may permit 20 to 25 institutions to participate in a multicenter trial without exorbitant equipment expenses.

In addition to the equipment issue, staff training and calibration of measurement are significant concerns. It is incumbent on study coordinators of multicenter trials to develop ongoing procedures for training and recalibration and to ensure

that funding to support these activities is included in all grant proposals. Training strategies may include sessions wherein all participants travel to a central site for training or may involve the development of a training team funded to conduct site visits at each institution on a rotating basis for training and recalibration purposes. The latter strategy may ultimately prove to be more cost-effective. In addition, this strategy guards against systematic drift that may occur in measurement if training and recalibration occur at the same time for all institutions.

Once steps have been taken to ensure that all institutions have adequate equipment and training, data quality assurance procedures remain necessary. Simple procedures, like establishing parameters for data that are entered into a computer database, may assist in identifying gross errors of data entry. However, this procedure is inadequate in maintaining accurate data. For instance, the computer parameters for appropriate data entry for full-scale IQ scores might range from 50 to 150. Transposition of a score (e.g., 68 for 86) would not be detected by these parameters, but would represent significantly different clinical performance. Quality control of the database in a multicenter trial requires central review prior to data entry, as well as database review for data input errors. Failure to establish routines for data quality assurance jeopardizes all other aspects of the study. These concerns apply to data management in single institution studies but grow in importance in the multicenter trial.

Problems and Conflicts in Decision Making Concerning Research

While having a multitude of individual professionals involved in research provides many benefits, collaboration across institutions and disciplines may lead to administrative structures that are unwieldy and sometimes interfere with efficient decision making. This is perhaps the most significant difficulty encountered in collaborative group research. In some instances, difficulties may arise because procedures for consensus building and final responsibility for decision-making fail to be developed prior to initiation of the research project.

A second issue of concern arises when the internal politics of a cooperative group clash with the institutional politics and policies of participating group members. Investigators who participate in cooperative group research necessarily have two allegiances, one to their employing institution and one to the cooperative group. Sometimes single institutional studies are proposed that may interfere with a larger collaborative study. For instance, one institution may have a new treatment that is believed to reduce toxicity while maintaining efficacy in the treatment of children with acute leukemia. One toxicity expected to be reduced might be memory impairment. However, the institution may have agreed to participate in a cooperative trial involving a different treatment and clinical investigation of the effects of this treatment on verbal fluency. Participation in the institutional study may preclude the agreement to participate in the cooperative study, creating a conflict.

Compliance with Protocols at Different Sites

One of the major impediments to successful implementation of a protocol within a multicenter group is the degree to which all participating institutions can comply with data collection requirements. This compliance issue may occur at the

institutional level because of limited resources, either personnel, equipment, or both, or at the group level, where guidelines for quality control exceed the resources of individual institutions. For psychologists, this conflict may arise when decisions about specific tests to be administered or intervention procedures to be used need to be made. It has been jokingly stated that ten psychologists working together to determine a measurement strategy will develop a consensus protocol when nine die and the tenth walks out of the room with his or her protocol in hand. Clearly, multicenter research can be effectively conducted only when all participating members are committed to compromise in order to resolve individual institution and group conflicts.

Investigator Attrition and Change

Another pragmatic issue that may arise within collaborative research arrangements occurs when one or more key investigators or support staff leave the group due to illness or job change. This same problem applies to individual research, but is compounded by the frequency and number of changes that may occur within the multicenter setting. For this reason, a continual need for retraining, whether at the investigator or clerical level, occurs in collaborative research. This demand often is unanticipated at the time when a project is initiated, but over time staff changes may substantially interfere with standardized collection of data and administration of a treatment protocol. Successful investigations plan for staff attrition; build in necessary duplication of knowledge, responsibilities, and skills; and provide for ongoing training of new institutional staff. Within a multidisciplinary framework, this approach may be quite difficult. Replacing a data manager or clinic nurse simply may involve teaching an individual with basic skills the specific information required by the study. However, replacing a skilled psychometrician, particularly one with expertise in evaluating critically ill children, may prove to be a daunting task at both the institutional and group level. It is critical for investigators to anticipate such staff changes when they evaluate the level of sophistication demanded by the research protocol over time (see Chapter 11, this volume).

Managing Conflicts

Many cooperative groups have developed strategies to resolve conflicts that occur in the course of multisite research. For instance, to resolve conflict between a single institution's proposed research and the cooperative research group, the institution may be given the group's permission to conduct an independent study with the support of the group, on the condition that the results of the institutional study will be considered as pilot data for a later cooperative trial. In other cases, selected institutions may cooperate to expand the measurement of cognitive functioning beyond a core battery administered at all institutions in the cooperative group. These strategies permit advancement of both institutional and cooperative group objectives. In each of these cases, however, agreement must occur in advance regarding the timing of dissemination of study results. This issue is a critical one, because premature publication of limited findings could adversely affect the completion of a larger, more comprehensive cooperative trial. Most cooperative groups develop policies for reviewing and resolving potential conflicts between institutional and cooperative group goals. The issues that result from these conflicts are

often quite difficult to resolve once problems occur but not insurmountable, especially if the organization of the research collaboration is carefully planned prior to the initiation of a study.

Structures to Support Consensus Building

Cooperative research requires that there be a consensus-building process of decision making. However, when individuals from multiple institutions and disciplinary backgrounds attempt to reach consensus, the process sometimes can be quite unwieldy. For this reason, the structure of cooperative research efforts must include procedures that support consensus building and effective mechanisms for reaching a decision when a consensus cannot be reached. Most successful cooperative groups have developed such structures. For instance, research group guidelines typically permit individual investigators to initiate a research concept for consideration by the group. In some cases this concept may involve a significant and unique clinical trial comparing one treatment with another. For instance, a psychologist might propose a study to compare a relaxation training intervention with conscious sedation in alleviating distress in children undergoing lumbar punctures in the treatment of cancer. In other cases, the proposal might represent an ancillary component of an existing study, for instance, adding a neuropsychological component to an existing trial comparing two types of central nervous system treatment for brain tumors. In both cases, the implementation of the proposal will require a commitment of resources, patient time, institutional resources (e.g., IRB review and approval), and patients available for accrual. Because of these costs and the potential impact of any given proposal on other concurrent studies, review and prioritization of the project by members of the group, across disciplines, are essential. The proposal is circulated first to investigators with similar discipline backgrounds and then to other affiliated investigators who might participate in such a project. Peer review of the proposal then occurs and a decision to proceed with or abandon the proposal is made.

Once a proposal is initiated by a single investigator or subgroup of researchers and receives support to move forward in accord with the above procedure, it is then circulated to investigators from each of the participating institutions for both scientific and practical review. Issues related to individual financial, resource, and institutional politics may arise at this point. It is not unusual for a scientifically solid research plan to fail to advance beyond the proposal phase simply because it is perceived as overburdensome or costly at the level of participating institutions. It is critical for investigators in cooperative research groups to establish a framework for anticipating and resolving conflicts that may arise between the group's concept of science and the individual needs of participating institutions. Some cooperative groups have committed a portion of group resources to support ancillary studies and have developed formal mechanisms for ancillary study review (e.g., CSSCD). Such procedures are often included to encourage new ideas and vitality in the activities of the cooperative research group.

Successful research groups also facilitate their own cooperative activities by developing an administrative hierarchy that permits consensus building, while at the same time providing sufficient authority and structure to facilitate decision making. There are a number of models for cooperative group administration, but most include similar principles. For example, a single individual typically serves

as chair of the group, and this person is typically elected by a vote of participating institutions. The group often elects an executive committee and establishes a mechanism for rotation of membership on the executive committee over time. The group chair, working with the advice of the executive committee, then appoints study coordinators, forms working committees, and appoints committee chairs. Decisions about guiding concepts and critical study issues are made by consensus at the committee levels, subject to review by committee chairs, the executive committee, and ultimately the group chair. Such a structure not only permits participation at all levels of scientific endeavor by all individual members of the cooperative group, but also provides a structure for appeal and review when consensus cannot be reached.

Professional Issues

While many of us would like to think of our research activities as objective scientific endeavors, in reality personal goals and professional recognition are often strong driving forces behind our activities. The need to have one's ideas accepted, to receive recognition for contributions, and to promote professional perspectives often enter into decisions about how multicenter group research is conducted. Three such problems have emerged as potential pitfalls in multicenter collaborative research: (1) scientific and political leadership within the group; (2) specific responsibilities and requirements for study coordination; and (3) recognition of effort through inclusion on decision-making committees and authorship of resulting manuscripts. Failure to recognize and address these areas may ultimately result in failure of a collaborative group to achieve its scientific goals.

Leadership in Multisite Research

Successful research groups recognize that there are two kinds of leadership necessary for cooperative clinical research across general centers. First, group participants must identify individuals who can provide effective political leadership for the group. These individuals must be highly skilled at developing communication between participants, facilitating open discussion about both scientific and pragmatic issues, advocating for the needs of the group and for individual group members, and facilitating negotiation and compromise. They also must be viewed by group members as objective, thus permitting them to arbitrate conflicts in a manner that allows group research activities to proceed effectively.

The second kind of leadership necessary is scientific, and this necessarily rests with an investigator who has substantial experience and demonstrated success as a researcher in the content area of the multisite research. It is possible for a group to combine both political and scientific leadership within a select group of individuals, but this is neither essential nor always practical. However, successful group leadership requires individuals involved in administrative leadership to be aware of the many diverse perspectives and needs related to different disciplines involved in scientific research.

Psychologists working in a multidisciplinary, multicenter group must be keenly aware of the above leadership issues and the implications for their work. Unless the core question of the scientific trial is psychological, it is likely that the primary leadership will be provided by a pediatrician or pediatric subspecialist.

While there may be a strong personal and scientific commitment to working with pediatric leadership, significant differences in training, clinical perspectives, and research approach may create difficulties for the collaboration. For example, training in development and behavior is limited for physicians, and their understanding of multivariate designs and variability in psychological data also may be quite limited. A great deal of emphasis in many clinical trials is placed on a categorical clinical endpoint such as disease–no disease. The understanding of variation around a mean or the meaning of an *F*-statistic, essential components of psychological research, may be quite foreign to some physicians. On the other hand, training in double-blind methodology, use of survival analysis, and knowledge of medical treatment side effects is infrequently provided to psychologists in either graduate school or postdoctoral settings (Armstrong, 1995). These differences in perspectives, approach, and knowledge may result in very successful scientific investigations; they also may provide the foundation for potentially disruptive conflict between disciplines. Disciplinary leadership and overall group leadership that recognize these differences can provide opportunities for ongoing education for all disciplines and can facilitate communication between investigators that leads to a high-quality science.

Interdisciplinary Differences in Coordination and Conduct of Procedures

Another area for potential professional conflict lies in the expectations that different disciplines have for study coordination. Once again, differences in the type of data and procedures used by different professional disciplines may require different approaches to study coordination and the conduct of research. Consider the example of a study of treatment of children with medulloblastoma that may involve administration of a regimen of chemotherapeutic agents, followed by collection of laboratory data. The psychological component of this study may involve comparison of neurological findings with formal testing of psychological functioning. The administration of chemotherapeutic agents involves having a physician write an order, a pharmacist prepare the medications, and a nurse administer the medication on schedule. Blood work may be drawn and sent to a laboratory, with results recorded on a flow sheet. On the other hand, the psychologist must schedule time to test the child, determine that the testing conditions provide valid data, score the child's responses, make all necessary conversions, and report the data along with any contextual information. The degree of effort required by the coordinator of the medical information and the psychological information may be substantially different in both time and effort. This difference may be significantly associated with rates of compliance in data collection. In order to avoid conflict between coordinators of medical and psychological data, it is often quite helpful to clearly understand the procedures used by each study coordinator and the unique limitations of each component of the study. Such communication may, in fact, make data collection for both study coordinators more efficient and result in higher data quality.

Difficulties Concerning Authorship

Most investigators encounter difficulties related to authorship at some point in their research careers. Authorship order is a particular problem in multicenter

collaborative research, because many individuals make significant contributions to the development of a concept, drafting of a study and/or grant proposal, data collection and management, statistical analysis, and manuscript preparation. Specific criteria for authorship on manuscripts submitted to biomedical journals include authors being involved in the project from design until manuscript preparation and able to take public responsibility for the findings of the study (International Committee of Medical Journal Editors, 1994). How these criteria can be best applied to collaborative research projects is not clear. For studies spanning years or decades, the continuity requirement sometimes cannot be met by any investigator. Studies of limited populations may take years to complete, during which time investigators may change and relative contributions may vary. Even if these criteria are relaxed to account for significant intellectual contributions, it may be impossible to include all the authors who make the necessary contribution on a journal masthead.

Strategies to Manage Authorship. In our experience, it is very important that negotiations about authorship credit and order take place at the time of study initiation and to continue as changes take place across the life of the study. Scientific and ethical obligations vary somewhat across disciplines, and there may be situations when the ethical framework of investigators is different. It is incumbent on the psychologists working with a multicenter collaborative group to clearly inform all colleagues about their professional guidelines related to authorship credit (American Psychological Association, 1992) and to negotiate authorship issues in a way that is consistent with the ethical principles of all disciplines involved.

Different strategies have been used to address the issue of authorship credit in cooperative groups. One strategy has been to abandon the concept of individual authorship and give credit for the published work to the cooperative group. While this strategy lessens tension about authorship order, it offers few benefits to investigators when promotion and tenure decisions are made, or when scientific contributions are reviewed as part of a grant award process.

A second strategy has involved giving senior authorship to those individuals who carry the primary coordination responsibilities for the study and including other investigators who contribute data from participating institutions as junior authors. A recent study found that participation in cooperative group research is valued and rewarded in promotion and tenure decisions by chairs of departments of pediatrics in Canada, although primary weighting is given to the principal investigators of such trials (Davies, Langley, & Speert, 1996).

When a journal limits the number of offers permitted on a masthead, some groups decide to spread authorship out across several papers, particularly in the case of junior authors, so that all participants are included as an offer on at least some of the group publications. In all cases, all contributing authors are acknowledged in a footnote to the paper.

Another issue that arises in cooperative group research that is different from independent investigator-initiated research is the necessity for group investment in disseminating findings through publication. Data resulting from a cooperative group effort belong to the cooperative group, not to the investigator assigned to conduct the study. When negotiated authors fail to produce manuscripts within reasonable time limits, some groups exercise the option of assigning authorship on a paper to another senior author or group of authors. This policy is typically

clearly stated in the bylaws of the group and is enforced to ensure that group data are published in a timely fashion. Investigators participating in corporate group research should be keenly aware of such policies so as to reduce conflict at the time a manuscript is due.

FUTURE DIRECTIONS

While opportunities for multicenter and multidisciplinary research collaboration for psychologists are relatively new, it is likely that this model will increase in popularity and use in the years to come. In order for this growth in multidisciplinary research to yield maximum benefit, research grants and training models need to support these activities.

Grant Support for Multicenter Research

Funding for multi-center trials can occur as a result of specific administrative or legislative action by which dollars are specifically targeted to support multi-institutional research within a given disease area. However, proposals for multi-institutional collaboration outside these targeted areas may be viewed as overly costly and burdensome, resulting in low priorities for funding. Several factors contribute to these costs. Multi-institutional research projects often require redundant duplication of personnel and resources at each participating institution. When individual investigators submit grants, there is a clear separation of direct costs of research and institutional or indirect costs. In multi-institutional studies, indirect costs associated with subcontracting institutions are usually considered as direct costs of the overall grant budget (US Department of Health and Human Services, Public Health Service, 1997), resulting in a final direct cost for research that appears excessive and serves a barrier to funding.

Several evolving models to support multicenter research are now being developed. Mechanisms for funding core activities, sharing equipment resources across multiple sites, developing redundancy and training and skills, and maintaining a high level of data quality are all critical to the further development in this area. Several strategies have been developed by the Psychology Committee of the Pediatric Oncology Group. These include: (1) credentialing institutions and investigators to participate in multicenter trials; (2) applying for core support for study coordinator time to monitor study enrollment and review data quality; (3) funding investigator time on a per patient enrollment basis; and (4) developing mechanisms for sharing expensive equipment. This approach is expected to reduce cost, increase efficiency, and increase the probability of extramural funding for multi-center research trial.

Interdisciplinary Research Training

As we have both noted in previous articles (Armstrong, 1995; Drotar, 1993) there is a strong need for interdisciplinary research training for both psychologists and pediatricians. This may require changes in graduate and medical education, but will most likely result in training opportunities at the postgraduate level.

Training programs that develop a curriculum in both biological and behavioral issues of chronic and acute illness in children will best be suited to meet this need. Such training will necessarily include education about disease, pathophysiology, medical treatment options, treatment side effects, behavioral aspects of disease and treatment, and late effects, while interdisciplinary training may occur informally as psychologists supervising graduate students involve these students in multidisciplinary medical teams. Developing formal mechanisms to participate in such training is also necessary. These may involve having senior pediatricians and pediatric psychologists participate in lecture series on major illnesses of childhood and moderately advanced disease and treatment information in a seminar format. Graduate students can be encouraged to develop thesis and dissertation committees that can be cochaired by a psychologist and a pediatrician. Research mentoring of pediatric residents and fellows by psychologists, as well as psychology graduate students and fellows by pediatricians, also may serve as models for training. Formation of multidisciplinary research groups and journal clubs also may facilitate this process. With a common basic clinical and scientific foundation, physicians and psychologists then can undergo training in research design and methodology that maximize each profession's understanding of a complete research model.

Several models of interdisciplinary research training are possible. At our institution, students and postdoctoral fellows participate in weekly pediatric hematology–oncology clinical rounds and attend monthly multidisciplinary neuro-oncology conferences that involve the review of MRI and computed tomography scans, neuropathology slides, chemotherapy protocols, and serial neuropsychological test data. Pediatric residents and fellows complete monthly elective rotations with pediatric psychologists working in primary care settings, as well as in a pediatric neuropsychology program. Such training can result in individuals in each discipline acquiring a shared knowledge base and may result in research projects that are more integrated and scientifically well grounded. For example, pediatric psychologists may then ask research questions that appropriately address and include critical medical variables; pediatricians are able to conduct medical studies with a clearer understanding of the mechanisms and appropriate measures of behavioral factors. When such training models work most efficiently, it is not unusual for psychologists to knowledgeably raise questions about relationships between disease mechanisms, medications, or radiographic data and behavioral outcomes, or for pediatricians to propose specific measurement strategies for assessing toxicity associated with a particular clinical treatment. For instance, for a number of years, oncologists in the POG only infrequently asked about the effects of treatment on intelligence. Years of collaboration with pediatric psychologists have helped them to recognize and understand complex relationships between the type of treatment and neural development, the child's age, and specific neuropsychological functions (such as memory, attention, visual–motor integration). On the other hand, pediatric psychologists have learned to be careful in their selection of measurement tools, depending on the specific chemotherapeutic agents being used to treat children with specific disease.

Opportunities also can be provided for graduate students to participate in multisite research by allowing them to work with data sets that have been gathered as a part of these projects, develop manuscripts in collaboration with investigators at different sites, and participate in investigators' meetings in which procedures,

data analytic tools, and new research directions, including proposals, are discussed. Such participation gives students an invaluable real life experience with the methods, problems, and potential of multicenter research.

FINAL THOUGHTS

Multidisciplinary and multicenter collaborative research offers an exciting set of opportunities and challenges for research in pediatric and child clinical psychology. Ultimately, the sophistication of research that can be conducted using this model will increase, and specialized training for investigators will be necessary. The research conducted by pediatricians, psychologists, clinical psychologists, and other professions offer a great deal to the health and well-being of children. Combining professional skills and expertise in sound scientific research provides the opportunity for significant advances in the quality of life of pediatric patients and their families.

REFERENCES

- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*.
- Armstrong, F. D. (1995). Commentary: Childhood cancer. *Journal of Pediatric Psychology*, 4, 417-421.
- Armstrong, F. D., Toledoano, S. R., Miloslovich, K., Lackman-Zeman, L., Levy, J. D., Gay, C. L., Schuman, W. B., & Fishkin, P. (in press). The Miami Pediatric Quality of Life Questionnaire: Parent Scale. *International Journal of Cancer*.
- Brocklehurst, P., Elbourne, D., Garcia, J., & McCandlish, R. (1996). Clinical trials and rare diseases: Trials of adequate size are possible with the right organization. *British Medical Journal*, 312, 910.
- Brown, R. T., Armstrong, F. D., & Eckman, J. R. (1993). Neurocognitive aspects of sickle cell disease. *Journal of Learning Disabilities*, 26, 33-45.
- Chordokoff, J., & Whitten, C. F. (1963). Intellectual status of children with sickle cell anemia. *Journal of Pediatrics*, 63, 29-35.
- Crooks, S. W., Colman, S. B., & Campbell, I. A. (1996). Costs and getting ethical approval deter doctors from participating in multicentre trials. *British Medical Journal*, 312, 1669.
- Davies, H. D., Langley, J. M., & Speert, D. P. (1996). Rating authors' contributions to collaborative research: The PICNIC survey of university departments of pediatrics. *Canadian Medical Association Journal*, 155, 877-882.
- Diabetes Control and Complications Trial Research Group. (1993). The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *New England Journal of Health*, 329, 977-986.
- Drotar, D. (1993). Influences on collaborative activities among psychologists and pediatricians: Implications for practice, training, and research. *Journal of Pediatric Psychology*, 18, 159-172.
- Drotar, D. (1995). *Consulting with pediatricians: Psychological perspectives*. New York: Plenum Press.
- Drotar, D., Agle, D. P., Eckl, C. L., & Thompson, P. A. (1996). Impact of the repressive personality style on the measurement of psychological distress in children and adolescents with chronic illness: An example from hemophilia. *Journal of Pediatric Psychology*, 21, 283-293.
- Farber, M. D., Koshy, M., & Kinney, T. R. (1985). Cooperative Study of Sickle Cell Disease: Demographic and socioeconomic characteristics of patients and families with sickle cell disease. *Journal of Chronic Diseases*, 38, 495-505.
- Fowler, M. G., Whitt, J. K., Lallanger, R. R., Nash, K. B., Atkinson, S. S., Wells, R. J., & McMillan, C. (1988). Neuropsychologic and academic functioning of children with sickle cell anemia. *Journal of Developmental and Behavioral Pediatrics*, 9, 213-220.
- Gaston, M., Smith, J., Gallagher, D., Flournoy-Gill, Z., West, S., Bellevue, R., Farber, M., Grover, R., Koshy, M., Ritchev, A. K., et al. (1987). Recruitment in the Cooperative Study of Sickle Cell Disease (CSSCD). *Controlled Clinical Trials*, 8, 131S-140S.

- Gonzalez, J. C., Routh, D. K., Saab, P. E., Armstrong, F. D., Guerra, E., Shifman, L., & Fawcett, N. (1989). Effects of parent presence on children's reactions to injections: Behavioral, physiological, and subjective aspects. *Journal of Pediatric Psychology*, 14, 449-462.
- Hardy, M. S., Armstrong, F. D., Routh, D. K., Albrecht, J., & Davis, J. (1994). Coping and communication among parents and children with human immunodeficiency virus and cancer. *Journal of Developmental and Behavioral Pediatrics*, 15, S49-S53.
- Hillgartner, M. W., Donfield, S. M., Willoughby, A., Contant, C. F. Jr., Evatt, B. L., Gomperts, E. D., Hoots, W. K., Jason, J., Loveland, K. A., McKinlay, S. M., & Stehbens, J. A. (1993). Hemophilia growth and development study: Design, methods and entry data. *American Journal of Pediatric Hematology Oncology*, 15, 208-218.
- International Committee of Medical Journal Editors. (1994). *Uniform requirements for manuscripts submitted to biomedical journals*. Philadelphia: American College of Physicians.
- Kazak, A. E., Penati, B., Boyer, B. A., Himelstein, B., Brophy, P., Waibel, M. K., Blackall, G. F., Daller, R., & Johnson, K. (1996). A randomized controlled prospective outcome study of a psychological and pharmacological intervention protocol for procedural distress in pediatric leukemia. *Journal of Pediatric Psychology*, 21, 615-632.
- Lantz, J., Reed, R. C., & Lewkowitz, S. (1994). Collaborative research-academic appointments: The myth and reality. *Nursing Administration Quarterly*, 18, 38-43.
- Lemanek, K. L., Horwitz, W., & Ohene-Frempong, K. (1994). A multiperspective investigation of social competence in children with sickle cell disease. *Journal of Pediatric Psychology*, 19, 443-456.
- Loveland, K. A., Stehbens, J., Contant, C., Bordeaux, J. D., Sirois, P., Bell, T. S., Hill, S., Scott, A., Bowman, M., Schiller, M., Watkins, J., Olson, R., Moylan, P., Cool, V., & Belden, B. (1994). Hemophilia growth and development study: Baseline neurodevelopmental findings. *Journal of Pediatric Psychology*, 19, 223-239.
- Morrow, G. R., & Bellig, A. J. (1994). Behavioral science in translational research and cancer control. *Cancer*, 74, 1409-1417.
- Onufrik, B., Saylor, C. F., Taylor, M. J., Eyberg, S. M., & Boyce, G. C. (1995). Determinants of responsiveness in mothers of children with intraventricular hemorrhage. *Journal of Pediatric Psychology*, 20, 587-599.
- Pediatric Oncology Group. (1992). Progress against childhood cancer: The Pediatric Oncology Group experience. *Pediatrics*, 89, 597-600.
- Pegelow, C. H., Armstrong, F. D., Light, S., Toledano, S. R., & Davis, J. (1991). Experience with the use of prophylactic penicillin in children with sickle cell anemia. *Journal of Pediatrics*, 118, 736-738.
- Quittner, A. L., Tolbert, V. E., Regoli, M. J., Orenstein, D. M., Hollingsworth, J. L., & Eigen, H. (1996). Development of the Role-play Inventory of Situations and Coping Strategies for parents of children with cystic fibrosis. *Journal of Pediatric Psychology*, 21, 209-235.
- Rosendorf, L. L., Dafni, U., Amato, D. A., Lunghofer, B., Bartlett, J. G., Leedom, J. M., Wara, D. W., Armstrong, J. A., Godfrey, E., Sukkestad, E., & Cunts, G. W. (1993). Performance evaluation in multicenter clinical trials: Development of a model by the AIDS Clinical Trials Group. *Controlled Clinical Trials*, 14, 523-537.
- Sargent, J. R., Sahler, O. J. Z., Roghmann, K. J., Mulhern, R. K., Barbarian, O. A., Carpenter, P. J., Copeland, D. R., Dolgin, M. J., & Zeltzer, L. K. (1995). Sibling Adaptation of Childhood Cancer Study: Siblings' perceptions of the cancer experience. *Journal of Pediatric Psychology*, 20, 151-164.
- Smyth, J. F., Mossman, J., Hall, R., Hepburn, S., Pinkerton, R., Richards, M., Thatcher, N., & Box, J. (1994). Conducting clinical research in the new NHS: The model of cancer. *British Medical Journal*, 309, 457-461.
- Snyder, C. R., Hoza, B., Pelham, W. E., Rapoff, M., Ware, L., Danovsky, Highberger, L., Rubinstein, H., & Stahl, K. J. (1997). *Journal of Pediatric Psychology*, 22, 399-421.
- Stowe, S. M., Hammond, G. D., Chard, R., Hornberger, J. A., Honour, R. C., Spoto, R., & Weiner, J. M. (1984). Monitoring clinical research. A report from the Children's Cancer Study Group. *American Journal of Clinical Oncology*, 7, 557-566.
- Swift, A. V., Cohen, M. J., Hynd, G. W., Wisenbaker, J. M., McKie, K. M., Makari, G., & McKie, V. (1989). Neuropsychological impairment in children with sickle cell anemia. *Pediatrics*, 84, 1077-1085.
- Thomas, A. M., Peterson, L., & Goldstein, D. (1997). Problem solving and diabetes regimen adherence by children and adolescents with IDDM in social pressure situations: A reflection of normal development. *Journal of Pediatric Psychology*, 22, 541-561.
- US Department of Health and Human Services, Public Health Service. (1997). Grant Application (PHS 398). OMB No. 0925-0001.

14

Ethical Issues in Conducting Research with Pediatric and Clinical Child Populations in Applied Settings

**DENNIS DROTAR, JAMES C. OVERHOLSER,
RACHEL LEVI, NATALIE WALDERS,
JANE R. ROBINSON, TONYA M. PALERMO
and KRISTIN A. RIEKERT**

Over and beyond the general ethical principles with which all psychologists should be thoroughly familiar (American Psychological Association, 1992; Canter, Bennett, Jones, & Nagy, 1994), research with pediatric and clinical child populations raises special ethical problems that challenge investigators to develop special strategies in order to conduct their research in a responsible manner. The work described in this chapter, which is based on research experiences in hospital, school, and clinic settings, describes a number of knotty ethical problems that have not been well articulated in previous work. These issues include considerations in working with groups that are charged with the oversight of research ethics in clinical settings, problems related to confidentiality of data, obtaining appropriate consent, managing risk-related to psychological vulnerability, for example, depression in adolescents, and maintaining appropriate role boundaries as researchers in clinical settings. Using illustrations from research with varied populations of children and families, we describe a number of ethical issues that have

DENNIS DROTAR, JANE R. ROBINSON, and TONYA M. PALERMO • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038. **JAMES C. OVERHOLSER, RACHEL LEVI, NATALIE WALDERS, and KRISTIN A. RIEKERT** • Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106-7123.

arisen in data collection in clinical settings and consider strategies to manage them. Our aim is to help researchers to anticipate and prevent ethical problems wherever possible and to provide suggestions concerning how their impact on families can be minimized.

Readers should recognize that there are many and varied ethical challenges involved in research with different populations of children and adolescents that are beyond the scope of this chapter. Consequently, interested readers might wish to consult other descriptions of specific ethical issues. For example, general ethical issues in research with children are described by Glantz (1996); Grodin and Glantz (1994); and Thompson (1990). Sources that consider research including informed consent issues with children and families include Abramovitch, Freedman, Henry, and VanBrunschot (1995), Fisher (1993), Keith-Spiegel and Koocher (1990), Koocher and Keith-Spiegel (1998), Melton, Koocher, and Saks (1983), Range and Cotton (1995), and Weithorn and Schearer (1994). Children's rights to privacy in research, including research with adolescents, are considered by Brooks-Gunn and Rotheram-Borus (1994), Keith-Spiegel and Koocher (1990), and Koocher and Keith-Spiegel (1998). Finally, we also refer you to descriptions of specialized ethical issues in research with specific populations and methods including minorities (Scott-Jones, 1994), children with acquired immunodeficiency syndrome (AIDS) (Rotheram-Borus & Koopman, 1994), child abuse (Hoagwood, 1994; Liss, 1994), family observational methods (Bussell, 1994), intervention research with high-risk children and youth (Fisher, 1993, 1994; Scarr, 1994), and research with children and adolescents with mental health problems (Hoagwood, Jensen, & Fisher, 1996).

WORKING WITH GROUPS WHO OVERSEE RESEARCH ETHICS IN CLINICAL SETTINGS

Clinical researchers need to be mindful of the rules that govern the conduct of research in specific clinical and applied settings. Consequently, one of the researcher's first tasks is to identify and work closely with individuals and/or groups who are responsible for monitoring the ethical standards of research conducted in particular settings. In our experience, it is much more effective to anticipate potential ethical issues and implement appropriate safeguards prior to data collection than it is to manage them after the study has begun. Investigators who do not take such a proactive approach run a greater risk of incurring significant ethical problems as well as objections from oversight groups and collaborating staff that could threaten the integrity of their data collection.

Considerations in Working with Institutional Review Boards

The majority of hospital settings and other institutional settings, for example, universities, have developed formal procedures for evaluating the ethical merits and problems associated with research. In many settings the primary oversight group is an institutional review board (IRB), which is an interdisciplinary group of professionals, for example, physicians, lawyers, and members of the lay public, who are charged with insuring that research procedures in their setting follow ethical standards. The National Research Act of 1974, which created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral

Research, mandated the establishment of IRBs in all research organizations receiving research funds, including research with human subjects (Department of Health and Human Services, 1983; Gray, 1977; Novick, 1981; Phillips, 1996).

Although IRB review procedures were established to facilitate the ethical conduct of medical research (Phillips, 1996), they also are very important to the clinical conduct of psychological research (Ceci, Peters, & Plotkin, 1985) [See Grodin and Glantz (1994) for a useful description of regulations and functions of IRBs.] Consequently, researchers should develop a close working relationship with representatives from their local IRB and become thoroughly familiar with local IRB rules, regulations, and operating procedures. Prior to preparing their proposals, investigators may wish to obtain consultation from IRB representatives concerning potential ethical concerns about their research and information concerning the criteria for expedited review procedures that involve no more than minimal risk to participants, which may be appropriate for some research projects (Grodin & Glantz, 1994). Wherever possible, psychologists also should take an active role in educating IRB members in their settings concerning the nature of psychological research. Becoming a member of the IRB is an excellent way to educate colleagues concerning the nature of psychological research and the special ethical issues that are involved in this work.

Researchers need to recognize that research reviewed by an IRB has to satisfy several critical requirements, including: (1) minimization of risk to subjects; (2) assurance that risks to subjects are reasonable in relation to anticipated benefits; (3) equitable selection of subjects; (4) appropriate informed consent procedures and documentation; (5) adequate provisions to monitoring data to ensure safety of subjects; (6) adequate provisions to protect the privacy of subjects and maintain confidentiality of data; and (7) implementation of special safeguards to protect the rights of subjects who are likely to be vulnerable to coercion and/or undue influence, for example, children, mentally disabled persons, and so on (Grodin & Glantz, 1994). Experienced researchers also learn to appreciate that different IRBs develop their own working guidelines for evaluation of protocols, which can include methodological rigor of study designs. In some settings, poorly designed studies can be viewed by IRBs as unethical because they waste participants' time and provide little benefit to society (Jonsen, 1978; Rutstein, 1969; Steiner, 1972). Investigators need to become familiar with the internal workings and guidelines that are utilized by their own IRBs in order to anticipate special concerns that may arise.

CONFIDENTIALITY ISSUES IN RESEARCH WITH CHILDREN AND ADOLESCENTS IN CLINICAL SETTINGS

A number of difficult problems related to confidentiality can arise in research with children and adolescents in clinical settings. These issues are considered in the next section.

Data Collection in Hospital and Clinical Settings

Confidentiality of data collection can be difficult to preserve in busy medical settings where space is at a premium. For example, in one clinical setting many children with human immunodeficiency virus (HIV) infection received intra-

venous treatments that required them to stay in the treatment room for several hours. While data collection in this setting provided an opportunity to conveniently collect interview data without interrupting the family's schedule, the presence of other families in the room and periodic interventions by the nursing staff threatened family privacy. As shown by this example, researchers who gather data in medical settings should anticipate the need to modify data collection procedures to enhance confidentiality, for example, using rooms that are less busy, speaking with nurses and physicians about giving some warning when they are coming, and so forth.

Other problems with confidentiality may be raised by situations where the staff physicians or nurses give the investigators additional information about a family and/or child that is unrelated to research. Such communications may arise from staff concerns about families that they are seeing. In the interests of courtesy, a researcher cannot easily walk away from such conversations as soon as they start. Nevertheless, we have found that it is very important that investigators limit these conversations by not reinforcing them, reminding the staff of their role as researchers, and redirecting these concerns to relevant clinical staff.

Special Issues in Confidentiality in Working with Children and Adolescents in Psychiatric Hospitals

Because of the very sensitive nature of many psychological studies with children and adolescents, it is essential that all participants be guaranteed confidentiality for all their responses. Any unnecessary and unauthorized breach of confidentiality could leave a researcher open to legal action (Dekraai & Sales, 1991). Consequently, researchers who work with children and adolescents in clinical settings, including hospitals, need to develop specialized training for their research staff concerning confidentiality issues. For example, in our research with adolescents who were admitted to a psychiatric hospital, our research assistants and staff receive training in the ethical standards related to the conduct of research on that setting. They then sign a confidentiality agreement form to ensure that they are familiar with the legal and ethical standards for research confidentiality (Appendix A) (Overholser, Adams, Lehnert, & Brickman, 1995a).

We also would recommend that researchers who conduct research in clinical settings develop a detailed record of the kinds of problems in confidentiality that can arise in their setting. Such information can be used to train staff who collect data and to develop strategies to limit these problems.

Trade-offs between Confidentiality and Risk to Participants

When research examines such topics as depression and suicide risk in high school students or discharged psychiatric patients, the trade-off between the need to maintain confidentiality versus the ethical need to respond to the potential risk associated with suicidality raises special problems. For example, in order to limit problems with confidentiality when assessing high school students, we have used anonymous data collection procedures. To accomplish this, when we collected data on level of depression and suicidality in a sample of high school students (Overholser et al., 1995a), consent forms were collected at the same time as

the research questionnaires but were immediately separated to prevent the possibility of identifying subjects' identities.

Although the above procedure ensured confidentiality, it prevented us from responding effectively if a subject reported high levels of depression or suicidal ideation. In order to anticipate the possibility that some research participants may report high levels of emotional distress as a cry for help, we forewarned subjects in the consent form that all information would be collected in this anonymous manner. Also, we provided the telephone numbers of our psychology department and the local suicide hotline in the consent form in case any subjects wanted to discuss their emotional distress with a trained professional. In this way, subjects knew that they should not expect a telephone call from the research team but were welcomed to call us or other mental health professionals. There is evidence to suggest that this type of nonintrusive design helps encourage more open and honest answers from subjects who are asked to report on their levels of emotional distress (Stanton, Burker, & Kershaw, 1991).

ETHICAL PROBLEMS IN RESEARCH WITH VULNERABLE CHILD AND ADOLESCENT POPULATIONS IN CLINICAL SETTINGS

Researchers who work with children and adolescents in clinical settings need to anticipate that there may be situations in which they may uncover problems (e.g., abuse, high levels of psychological distress) in their subjects that may require them to implement special safeguards. Some of these issues are considered in the next section.

Dealing with Research Participants Who May Be Depressed

Our experience in conducting research with pediatric populations (e.g., mothers of pediatric patients and/or adolescents who are depressed) suggests that safeguards for depressed research participants also may need to be implemented in pediatric settings. For example, in conducting research with mothers of infants who were failing to thrive or with adolescents who were pregnant, we have notified subjects who reported high levels of depression or distress on psychological measures of this fact. These research participants were then informed about available treatment resources that could help them with their psychological distress if they so chose (Robinson, 1998). Whenever research participants and/or providers who are involved in the child's care are notified about clinically significant psychological distress as defined by their response to a research protocol, these procedures should be carefully described in the consent form and discussed with potential participants.

Researchers who work in applied settings with pediatric and clinical child populations should also understand that their applications of what appear to be straightforward safeguards for their research participants may have unintended results, as shown in the following example. Ms. Johnson was a teenage mother of a 9-month-old girl diagnosed with failure to thrive. The mother demonstrated a high level of depression on the Beck Depression Inventory and also experienced intrusive thoughts regarding suicide. The investigator informed her about treat-

ment resources, and, as was described in the consent form, told her she would share this information with her daughter's pediatrician. At the time, she found this discussion to be useful and stated she wanted to discuss her feelings further. Consequently, with her agreement she was referred to the clinic social worker for additional evaluation. However, at a subsequent visit with the social worker, this mother denied any suicidal thoughts, and claimed to be experiencing no depression and to have no idea what the social worker was talking about. She then missed the next four follow-up appointments for her daughter. When she finally did return to the clinic, her child had lost weight. We eventually learned that the mother thought her child would be taken away from her if she said she was depressed. Fearing this outcome, she did not return for appointments. After several consultations with the social worker, this mother eventually agreed about her need for mental health services and obtained treatment for herself. Her child remains with her.

Managing Emergency Situations with Research Participants: The Severely Depressed or Suicidal Adolescent

Some research projects in clinical settings require investigators to anticipate emergency situations that could occur with their child or adolescent research participants and develop strategies to manage them (Brinkman, Overholser, and Klier, 1994). For example, when conducting research on adolescent depression, it is important to plan for the identification of subjects who are severely depressed or potentially suicidal (Burbach, Farha, & Thorpe, 1986). Investigators who conduct such research also should have a plan for how to intervene should a subject report potentially dangerous levels of depression or suicide risk. In such situations, subjects should be informed about the limits of confidentiality in the consent form, which may include the investigator's sharing of information with the subject's treatment team.

While our research experiences with adolescents who have been hospitalized for depression in several studies have indicated that emergency situations rarely arise, nevertheless we have found that it is wise to prepare written procedures to follow in the case of an emergency. For example, in our research on follow-up assessment of adolescent suicide attempters (Brinkman-Sull, Overholser, and Silverman, 1997), we were concerned about the risks involved when evaluating depression severity and suicide risk in an outpatient sample. In order to anticipate and manage such risks in the event that they occur, we have our investigators forewarn subjects about possible limits of confidentiality, evaluate the severity of depression and suicide risk, and be prepared to help the subject manage their situation (see Appendix B for more information). We recommend that other investigators develop similar fail-safe plans.

INFORMED CONSENT ISSUES IN RESEARCH WITH PEDIATRIC AND CHILD CLINICAL POPULATIONS

Researchers who work with children in clinical settings need to be prepared to develop and use understandable consent forms and anticipate and manage any special ethical constraints on informed consent that arise.

Special Considerations in Obtaining Informed Consent from Families

When working with children or adolescents, the investigator must obtain written consent from the child's parents or guardians, except when working with emancipated minors (Dekraai & Sales, 1991). However, the procedure to obtain consent from the child and parents can be very complicated when the child is brought for treatment by a custodial grandparent or when parents are divorced and have joint custody of the child. In such circumstances, researchers need to establish who is the child's legal guardian before obtaining consent.

When collecting data in medical settings, to ensure the family's confidentiality, the investigator should work closely with the child's attending physician in implementing the research and decide together on the best strategy to inform families about the research. In some settings, physicians may want to discuss the study with families directly. In others, a letter to eligible families from the physician and psychologist researcher may be the best way to inform families about the study. Irrespective of how families are informed about the research, the decision concerning the optimal method of contact needs to be made in collaboration with the professional, for example, physician, psychologist, and so forth, who has the primary clinical relationship with the child and/or family.

Prepare Understandable Consent Forms for Children and Families

The preparation of a clear, comprehensive, and informative consent form to inform children and families about the nature of their research is an important task for researchers who work with children and families. Grodin and Glantz's (1994) description of the key elements of informed consent include: (1) a clear statement that the study involves research; (2) an explanation of the purposes of research; (3) the expected duration of the participants' involvement in procedures; and (4) clear identification of any procedures that are experimental. In addition, consent forms should include information about the research procedures, reasonably foreseeable risks, potential benefits of the research to subjects or others (if relevant), alternative procedures, such as, other treatments that are available in a treatment outcome study, confidentiality of research participation, a clear statement that participation is voluntary and can be terminated at any point, and an explanation of whom to contact for answers to questions about procedures and problems that are encountered (Roth & Appelbaum, 1983). Research participants should be given a copy of the consent form if they want one. For this reason, it may be useful for researchers to carry a separate pack of extra consent forms to make it easy to provide a copy to the participants.

In order to be most user-friendly, consent forms should be as brief as possible, written in conversational language, and focused on the most relevant issues of the research. In our experience, it can be difficult to design the consent forms so that they can be understood by parents and children from a range of educational backgrounds. In fact, some studies have found that consent forms were almost as difficult to read as medical journals, requiring at least a college-level reading ability in order to comprehend the key sections (Morrow, 1980; Ogloff & Otto, 1991). In order to develop consent forms that can be read and understood by participants, researchers may need to obtain pilot data concerning the consent form, including readability and ease of comprehension.

In some projects, in order to facilitate child and family members' comprehension of the nature and expectations of research, it may be most effective for the investigator to read and explain the consent form to all subjects. This allows the investigator an opportunity to explain and clarify any technical passages and provides the subject an opportunity to ask questions about the study (Appelbaum & Roth, 1982).

Anticipating and Managing Special Constraints on Informed Consent

The ethical requirements of informed consent with children and families place extraordinary demands on investigators to consider any special considerations (e.g., cultural, educational, and literacy backgrounds) that impose constraints on consent and to develop effective strategies to anticipate and manage these problems. We have identified several predictable sources of constraint on informed consent that have occurred in research with children and families in clinical settings. These include families who consent to participate in research without a careful reading, problems in literacy, and constraints imposed by the behavior of professional staff.

Families Who Agree to Participate without Carefully Reading the Consent Form

One of the investigator's primary ethical responsibilities is to ensure that parents and children read and understand the informed consent form. This seemingly straightforward responsibility can be difficult to implement. For example, in the course of conducting various research projects in clinical settings, we have encountered some parents who listen to the description of the project, glance at the consent form without reading it, and say, "That's fine, where do I sign?" Because investigators have a vested interest in enrolling as many participants as possible in their studies, it can be tempting for them to accept such signatures as indicating informed consent. However, research ethics require investigators to make sure that parents carefully read the consent forms; to discuss their studies with parents in detail and have parents sign the form only after it is clear that they understand what is asked of them. To reinforce these safeguards, we have found it to be useful to verify the parents' understanding of their child's participation in the study by taking time to have them describe the study after they have read the consent form.

Anticipating the Impact of Parental Literacy Level on the Consent Process

In conducting research projects with various pediatric and child clinical populations, we have been concerned that parents' reading ability may limit their capacity to understand the consent document. For example, in a study of screening for maternal depression in a clinic setting, we found that many mothers whose children received care in this setting had limited reading and comprehension skills. Once having identified skill problems, we also were impressed with how difficult it was for many parents to admit that they had a problem with reading, even when they were asked directly about this issue.

For the above reasons, we recommend that investigators who work with pediatric and clinical child populations that might be expected to have difficulties with literacy need to anticipate the impact of this problem, not only on the validity of informed consent, but on the administration of measures. Pilot work provides an excellent way to obtain information. If at all feasible, it may be best to administer all informed consent and procedures in a standardized interview format so that problems with parents' and/or children's literacy levels are avoided.

Constraints on Informed Consent Imposed by Others

Another interesting set of "lessons from the field" concerning the ethical application of informed consent in clinical settings concerns the constraints on informed consent that can be imposed by investigators and their colleagues. While it is important that co-investigators communicate enthusiasm about the need for a parent and/or child to participate in a study, in some cases, such enthusiasm can operate as a constraint on informed consent. For example, collaborating physicians and other staff may use their authority to tell families: "I want you to participate in this study," rather than presenting it as a true option. Unless they are working closely with colleagues who present the study to families, it may be difficult for investigators to detect such problems, especially in settings where many staff interact with families.

A good example of the above problem occurred in the course of our research on intervention with families of infants with failure to thrive (Drotar & Robinson, 1998). Early in the course of this study, we learned that some of the pediatric residents told parents whose children were hospitalized for failure to thrive that if they did not participate in the intervention study they could be reported to the Department of Human Services. We identified this obvious violation of informed consent only because our project coordinator happened to overhear such a conversation with a family. Fortunately, this incident occurred sufficiently early in the course of the study that we could take the following corrective measures: We discussed this problem with resident and faculty physicians in the setting and underscored the critical necessity of using an appropriate informed consent procedure. In addition, we instructed our staff who presented the study to families to make sure that families understood that their participation was an option. We also gave a sample script to physicians that described our informed consent procedure, including our conversation with families about their options. Finally, we instructed our staff to slow the consent process down by asking families to take time to think about their decision before committing to the project.

This example indicates that investigators should anticipate that some practitioners in clinical settings may have a vested interest in having families participate in research projects. To limit, if not prevent such problems, it is important that researchers emphasize the importance of obtaining informed consent when introducing the study to practitioners in such settings.

Other problems in the consent process involve inconsistent styles of obtaining consent. For example, some investigators or research assistants may be more forceful than others in presenting the study to help families. To help ensure consistency and accuracy in the procedures for obtaining informed consent, we have found it helpful to develop standardized scripts to present information about

the study to families. All staff members who are in a position to introduce the study to families should have access to a standardized script that describes what to say to families about the study and should receive training in securing informed consent. Investigators also may wish to make sure that staff who present the study to families are observed in their interactions with families concerning consent.

Understanding Influences on Children and Families' Refusal to Consent

In the spirit of scientific inquiry, we recommend that investigators gather information concerning the reasons that potentially eligible research participants refuse their consent. Such information is important for several reasons. First, it is useful for investigators to identify the reasons for nonparticipation in research because this can facilitate understanding of sample selection bias (see Chapter 4, this volume). Moreover, researchers also need to identify potential barriers to research participation. Parents' reservations about participating in research may reflect potentially correctable misunderstanding. For this reason, when a parent or child refuses to participate in a study, it can be helpful to ask him or her why they do not want to participate and be sure they understand the nature of the study (Stuart, 1978). In some cases, a refusal to participate may reflect a misunderstanding about procedures that can be corrected through additional explanation.

SPECIAL ETHICAL ISSUES IN CLINICAL RESEARCH WITH CHILDREN

Special ethical concerns also may arise from specific types of psychological research conducted with children in clinical settings. Some of these problems are described in the next section.

Ethical Issues in Research Concerning Psychological Assessment

Research with children and adolescents can involve a wide range of specialized assessment procedures (see Chapter 5, this volume). For this reason, psychologists who work with children and families need to develop and maintain competence in specialized assessment procedures appropriate for these populations (Schwitzgebel, 1978). Ethically sound psychological assessment research requires up-to-date knowledge of the strengths and limits of a test, good clinical judgment, and skill in administering, scoring, and interpreting tests (Weiner, 1989).

Research with children and families requires specialized skills in engaging children and families in assessments. Consequently, it is important to consider the training and experience levels of individuals who collect data in pediatric and clinical child settings. For studies that involve specialized tests (e.g., developmental testing) and/or highly sensitive procedures (e.g., interviews with children concerning their psychological symptoms) (LaGreca, 1990), it may be best to use research assistants who have more advanced training, because they may be more prepared to give such procedures.

Conducting Treatment Studies

In treatment outcome research including studies of children, problematic ethical issues can arise in the selection of an appropriate control group. Ethical concerns may prohibit the use of a no-treatment control group for some populations, especially if treatment efficacy has been established. On the other hand, it is ethically defensible to compare a new treatment with standard care in the community for clinical problems for which the benefits of psychological treatment has not been established. Another option in treatment research is to compare the efficacy of two alternative treatments, each of which could be beneficial. For example, Schwartz, Chesney, Irvine, and Keefe (1997) have described the utility of comparing alternative intervention models that involve a similar level of contact with patients but very different assumptions about the underlying mechanism of therapeutic action and expectations for outcomes.

Other ethical considerations relate specifically to implementation of treatment studies in clinical settings. When a randomized control design is utilized, it is essential that children and families are randomly assigned to the different groups. However, such studies can raise ethical problems if treatment for certain conditions can be obtained only through the research protocol and if treatment is not available to nonparticipants (Lowe, Alexander, & Mishkin, 1974).

Special Ethical Considerations in Follow-up Studies

When research participation extends over time, as is often the case in longitudinal research with pediatric and clinical child populations, the investigator should not assume that families' initial consent implies a sustained willingness to participate in all subsequent phases of the study. On the contrary, consent needs to be reviewed at different phases of the research. Consequently, when planning a follow-up study, it can be important for investigators to obtain preliminary consent for subsequent contacts that would permit contacting the families at a later date to secure their consent for subsequent participation.

Another potential problem in clinical follow-up studies is the difficulty of maintaining confidentiality when multiple contacts, for example, phone calls, are necessary to locate former clinical patients for follow-up assessments in research (Nelson & Grunbaum, 1972). One strategy to manage this problem is to train research staff to contact the subjects and/or their families directly, ideally leaving no messages that could identify subjects' prior mental health treatment. For example, when the subject is not available, it may be better for researchers to ask for the subject by first name, and if asked to identify themselves, to provide a first name without stating the purpose.

Another strategy is to use a code number that disables caller identification when making calls to research participants. For example, in follow-up research of adolescents who had been admitted to psychiatric hospitals, once the initial assessment has been completed, we have asked the adolescent for permission to contact them for follow-up evaluation in 3 months (Overholser et al., 1995a). At the time of initial assessment, we ask adolescents to provide the best address and phone number, as well as a backup phone number of a relative or close friend whom we could contact to help locate them. At the time of initial consent, we then have

explained to adolescents and their families that their consent has allowed us to contact them later but does not require them to participate in the subsequent assessment. Also, when conducting follow-up assessments over the telephone, the preliminary written consent allows us to use a verbal consent at the time of follow-up. Finally, wherever it is feasible, we have found it helpful for the same investigator who conducted the initial assessment also to do the follow-up with a particular subject. This procedure can reduce subjects' concerns about confidentiality.

ENHANCING THE BENEFITS OF PSYCHOLOGICAL RESEARCH TO CHILDREN AND FAMILIES

In evaluating the ethical costs versus benefits of treatment and/or assessment research to participants, investigators need to address such questions as: "What is the benefit for the subject?" "Is there a potential for improving the quality of clinical care now or in the foreseeable future based on the research?" Ideally, wherever possible, psychological research should be integrated with clinical care and should facilitate the quality of treatment, education, and recovery for children and adolescents with different medical and physical disorders (for example, see Chapter 24, this volume). Consequently, investigators who work with pediatric and child clinical populations should work to provide certain direct benefits to subjects, where this is feasible. For example, in our research on depression among adolescent psychiatric inpatients (Overholser, Brinkman, Lehnert, & Ricciardi, 1995b), we obtained patient consent to provide feedback concerning standardized assessments to the treatment team to utilize in treatment planning. Our consent form (see Appendix C) includes permission to release information from our assessment to the adolescents' treatment team. We were able to provide the treatment team with a brief written summary of the child's test scores, along with both normative data reported by the test developers and normative data obtained from our research in the hospital and local schools. To facilitate the team's understanding of the test data, we also included a clear interpretation of cutoff scores.

Communicating with Other Professions about the Benefits and Limits of Psychological Research

While it is important to consider ways that research can be useful to practitioners in clinical settings, researchers also need to be careful about communicating misleading implications from research data that do not have demonstrated clinical validity (Messick, 1980). For example, when our research group conducted research on a new sense of humor scale that had established reliability or validity data (Freiheit, Overholser, & Lehnert, 1988), we felt it was not ethically defensible to provide feedback to the treatment team regarding the patient's scores on this new scale.

Researchers who work in medical settings also need to recognize that medical and nursing staff may give more credence to data from psychological research in clinical decision making than may be warranted (Drotar, 1989). Such problems are particularly likely to occur when physicians feel pressured to use research-based information to guide their decision making for difficult clinical problems. For

example, we have found that pediatricians may ask researchers for their clinical opinions about what to do with problems such as failure to thrive based on research data such as observations of parent–child interactions. In such instances, we have found that it is important to inform pediatric colleagues that such data are not clinically interpretable and that they also are confidential to the investigator (see subsequent section on maintaining the researcher's role).

On the other hand, while research data from individual patients should not be communicated to medical staff, medical and nursing colleagues appreciate summaries of group data that is based on patients in their practice. For example, pediatric subspecialists have been interested in data that we had gathered concerning the allocation of treatment-related responsibilities among parents, children, and adolescents for chronic conditions such as cystic fibrosis and diabetes. These specialists were somewhat surprised to learn that adolescents with these chronic health conditions were not as independent in managing their treatments as they had assumed, but still relied a great deal on their parents to help monitor their treatments and provide direct help (Drotar & Ivers, 1994).

Enhancing Understanding Parent and Child Perceptions of the Costs versus Benefits of Research

An important but ill-defined issue for researchers in pediatric and clinical child psychology concerns how parents and children perceive the costs versus benefits of their participation in psychological research (Stanley, Sieber, & Melton, 1987). In our experience, some parents may regard psychological studies as an unnecessary demand, an invasion of privacy, or even as stressful. Others may enjoy participating in research because it can provide an opportunity to share their reactions to stressful events (e.g., the diagnosis and treatment of a child with chronic illness), as shown by the following comment from a parent of a child with cancer: “People don’t always ask me how I feel about my son’s problem anymore and I appreciated the opportunity; it was helpful for me to be able to talk about it and have someone listen, even if it was research.” Other parents of children with chronic health conditions may appreciate the opportunity to obtain information concerning their child’s psychological status that is afforded by their participation in certain studies.

On the other hand, researchers also need to be cognizant of situations in which parents may have unrealistic expectations about the benefits their children will receive through their participation in psychological research. The fact that consent forms typically contain information about the absence of direct benefits of research does not preclude parents from having unrealistic expectations about the benefits of their participation. For example, a mother of a child with a chronic illness who gave her consent for her child to participate in a study of quality of life was enthusiastic about her child’s participation because she felt that she could “finally get some answers about how her child was really feeling about his illness.” She needed to be informed about the specific feedback that she would and would not receive if she participated.

Researchers should be alert to the fact that some parents and children also may experience unexpected distress by participating in psychological research. Clear communication with parents during the process of informed consent can provide

useful data concerning parents' concerns about their child's participation in research. For example, one mother of a child with cancer was initially interested in participating in a study of children's quality of life as perceived by mothers and children. However, once she understood that her son needed to participate in the research as well, this mother said that she just could not allow herself to discuss the study with him. She felt that she could not ask her son to participate because she did not want him to have to think about the quality of his life, given how much he had been through in his short life.

MAINTAINING THE ROLE OF RESEARCHER IN CLINICAL SETTINGS

Requests for Support and Clinical Intervention

One of the most consistent ethical dilemmas that we have encountered in conducting research in clinical settings involves maintaining clear and appropriate boundaries between the role of researcher versus clinician. In our experience, practitioners in high-pressure, high-need clinical settings may be tempted to construe the psychologist's role in their setting as a practitioner, not as researcher. Consequently, psychologists and other researchers who gather data in clinical settings should anticipate that they may be asked by parents, children, or staff to function in clinical roles.

The management of these requests poses significant problems: On the one hand, the ethical mandate to avoid blurring the roles of researcher and practitioner is quite clear. On the other hand, to abruptly dismiss the staff's clinical concerns with a curt: "I can't help you. I'm a researcher," could be easily seen as withholding and/or insensitive. For this reason, researchers need to listen to the concerns that the staff expresses, to empathize with these concerns, and to help staff determine to whom they can refer patients. For example, in the course of data collection for a study of quality of life and childhood chronic illness, a researcher was asked for advice by the nurse practitioner concerning the management of a clinical problem. A nurse received an alarming call from a mother who informed her about sexual exploration that her child had been engaging in with a cousin and wanted to know if the investigator would talk to the mother on the phone. The researcher did not have clinical responsibility in this setting and the nurse was redirected to make a referral to a staff social worker.

Requests for therapy and clinical intervention that are made by families in the context of research also need to be anticipated. For example, in a study of the psychological adjustment of children with chronic illness, one parent stated that she looked forward to asking the investigator's advice about how to manage her behavior. The parent was reprimanded of the purpose of the study; that she would not have an opportunity to ask about management of her child's behaviors, but could take advantage of other resources in order to obtain such information.

Dealing with Questions and Concerns from Research Participants about Their Child's Medical-Psychiatric Treatment

Another potential dilemma in maintaining appropriate boundaries of the researcher's role occurs when researchers are asked by parents to field medical questions about their child's medical and/or psychiatric care. In such cases, it is

very important to redirect parents to ask their questions of their child's medical provider, even if the questions appear relatively straightforward. Other difficult concerns are raised by families who communicate misgivings about their child's medical care or providers to researcher as in the following example: "Dr. X doesn't understand my child. During the last visit, she was mean to him. I don't know if my child is getting the best care. What do you think? Would you talk to her?" While the researcher's role clearly does not extend to answering such questions, empathizing with the parents' concerns and encouraging direct discussion with the physician are adaptive responses to such requests. Adolescent patients also may complain about the quality of their medical and psychiatric treatment to researchers. Such problems often can be managed simply by giving adolescents opportunities to express their feelings of dissatisfaction.

In our research projects we have found that some adolescents who are assessed for depression may ask the investigator about the nature of depression or its treatment. While it can be appropriate to provide background information about the symptoms of depression or common treatments, again it is critical for the investigator to retain the role of researcher. Consequently, it is important to answer such questions in a factual but general manner, informing subjects about general research findings that pertain to their questions. Participants' questions about their symptoms or treatment always should be referred back to the treatment team for more definitive responses.

Anticipating and Managing Problems in Maintaining the Researcher's Role

Researchers should anticipate that they will receive requests from staff to operate as a practitioner and discuss this problem with staff prior to the research. When they present their studies to staff in a particular setting, investigators can clarify the purpose of the study, their role in the study, and the staff's role with statements such as the following: "As a researcher I will not be able to help you directly with clinical problems that may arise in the course of this research. However, if you wish, I can certainly help you decide who is best to talk to about the problem. Moreover, after the study is completed, we also can discuss any implications of the findings of psychological needs of children in this setting and plan for services."

Another proactive strategy to limit practitioners' requests for researchers to function as practitioners in clinical settings is to ensure that there is adequate coverage for psychological services in settings where research is being conducted. Wherever possible, practitioners and researchers in pediatric settings should work together to develop integrated research and clinical programs. The findings from some psychological research projects can be used effectively to plan for psychological services and to inform the kind of care that is delivered (see Chapter 24, this volume).

FUTURE DIRECTIONS

Our description of the ethical problems that can arise in clinical settings with children and adolescents raises a number of important issues that need to be addressed in future research and training programs. These issues are considered in this final section.

Description of Ethical Issues that Arise in Different Settings and Populations

The present discussion of ethical considerations in conducting psychological research with pediatric and clinical child populations reflects the specific experiences of a relatively small group of investigators. While we believe that many of our experiences are generalizable to other settings and investigators, we also recognize that investigators in other settings may encounter other important ethical problems that need to be described and considered by other researchers. Consequently, we would strongly encourage additional published reports of ethical issues that arise from research in clinical settings and suggested solutions from researchers in pediatric and clinical child psychology. Moreover, difficult problems that are not adequately covered by the ethical principles (American Psychological Association, 1992) should be referred to the American Psychological Association's task force, which will revise the code of ethics for psychologists (Fischer, 1998).

Research on Ethical Issues in Pediatric and Clinical Child Populations and Strategies for their Solution

Our observations have indicated that we need to understand more about children's and parents' understanding of and reactions to participation in psychological research in a range of clinical settings. Important questions that are in need of data include the following: How do parents and children appraise the value, that is, costs and benefits of participating in psychological research? Do they perceive special constraints associated with their participation in research (Abramovitch et al., 1995; Stanley et al., 1987)? Recent empirical work that has documented parents' understanding of and reactions to participation in research concerning cancer treatment (Ruccione, Krimer, & Moore, 1991) and clinical research (Harth & Thong, 1995) underscore the utility of qualitative and survey approaches to assess participants' perceptions costs versus benefits of psychological research with children, adolescents, and their parents.

Stanley and co-workers' (1987) proposed agenda to inform researchers about ethical issues includes several key areas that are relevant to pediatric and clinical child populations: (1) communication between the researcher and participants, i.e., disclosure and comprehension of consent information reactions to informed consent procedures and decision making; (2) the competency of research participants including children (Weithorn & Campbell, 1982); and (3) the behavior of researchers, including factors that influence the development of ethical behavior in researchers. Researchers in pediatric and clinical child psychology have much to contribute to this agenda.

Training in Ethical Conduct of Research with Pediatric and Clinical Child Populations

Another important area of continuing need is training in ethics for graduate students and professional psychologists who conduct research with pediatric and clinical child populations. Research with children and adolescents in clinical settings raises many special problems that are not clearly described in ethical guidelines but have been identified by experienced researchers. Students need to be apprised of these problems and potential strategies for their solution. To accomplish this, a course on research ethics that focuses on general professional issues,

issues in publishing, and writing (Keith-Spiegel & Koocher, 1990), as well as ethical issues in research with children (Fisher, 1993, 1994; Grodin & Glantz, 1994) has been developed. Continuing education courses concerning salient ethical issues in conducting research with children provide another important opportunity for learning (Grodin & Glantz, 1994). One example of the necessity for such training is underscored by Range and Cotton (1995), who found that two fifths of published work in such journals as *Journal of Pediatric Psychology*, *Child Development*, and *Journal of Clinical Child Psychology* failed to specify parental permission or consent. Moreover, the majority of studies failed to specify assent from children (Range & Cotton, 1995).

Finally, investigators who conduct research with children in clinical settings should take special care to provide training for students and staff who gather data from these populations. Such training methods could involve role-playing of critical incidents in informed consent, issues in recruiting culturally diverse samples, identifying problems in literacy in family members, and developing strategies to understand and solve problems that are likely to limit informed consent. We encourage investigators to develop such training problems and report their experiences.

APPENDIX A

Student Confidentiality Agreement Form

As part of my involvement in psychological research conducted at Case Western Reserve University under the supervision of Dr. _____, I will be given access to confidential material regarding psychiatric patients and/or research subjects. I will be expected to consider all material strictly confidential. This means you will not discuss the material with anyone outside of the research setting.

According to American Psychological Association (APA) Standards, "Psychologists have a primary obligation to respect the confidentiality of information obtained from persons in the course of their work as psychologists. . . . Information obtained in clinical or consulting relationships, or evaluative data concerning children, students, employees, and others, is discussed only for professional purposes and only with persons clearly concerned with the case" (APA, Principle 5).

Psychologists should not discuss their clients with anyone who is not directly involved in the client's care. When the individual is a subject who was evaluated as part of a research study, information may be sometimes disclosed to the subject's treatment team. When this is done, the goal is to ensure that the client receives the best possible care. Only information that is essential for understanding the clients and their problems should be discussed. Also, you should not discuss information about specific subjects with other graduate students outside of the research setting. No information about subjects should be released to parties not directly involved in the client's care without their explicit, written permission unless the release of information is required by law or by a court order.

I agree not to discuss confidential material about research subjects outside of research supervision. Failure to comply with these guidelines may constitute adequate reason for termination from the research involvement.

Signed

Date

APPENDIX B

Emergency Procedures Guidelines for Follow-Up Evaluations

An adolescent will be considered at risk if:

1. He or she scores higher than 20 on the Children's Depression Index or higher than 6 on the Hopelessness Scale for Children.
2. He or she admits to active suicidal ideation on the suicide question of the Children's Depression Inventory.
3. He or she directly admits to active suicidal ideation during the follow-up assessment.
4. He or she appears to be emotionally distressed as judged by the investigator.

If any of the above are present the following emergency steps should be taken:

1. Let the adolescent know about your concern.
2. If the adolescent is under age 18, tell his parent or parents about your concern.
3. If the adolescent is 18 years old or older, try to obtain a signed release of information form from the subject that will allow you to tell his parent or parents about the subject's distress.
4. If the adolescent is currently in therapy, find out the name of his or her therapist and suggest that the subject discuss these feelings with the therapist. Also, try to obtain a signed release of information form that will allow you to speak with the subject's therapist. If a release form is obtained, inform the therapist of this information as soon as is possible.
5. If the adolescent is not currently in therapy but was in therapy at some time during the follow-up period, ask the subject for his or her previous therapist's name and suggest that the subject recontact this therapist to discuss his or her current feelings. Try to obtain a signed release of information form from the subject to allow you to speak with the subject's therapist. If a release form is obtained, inform the therapist of this information as soon as is possible.
6. If the adolescent has not had any follow-up treatment since his or her hospitalization, give the subject and the underage subject's parents a referral list with names and phone numbers of several appropriate psychologists or psychiatrists in the area. Urge them to follow through with your referral.
7. If the adolescent is felt to be in immediate risk of harming him- or herself and in need of immediate hospitalization, inform the subject and the underage subject's parent(s) that you feel hospitalization is needed and call a hospital staff psychiatrist who is on call for that evening.
8. If the adolescent describes a plan to hurt another person or describes the presence of child abuse by themselves or an adult, this information must be reported to the authorities.
9. Plans of violence to another should be reported to the underage subject's parent(s) and the police department within the subject's home community should be notified.

APPENDIX C**Sample Consent Form*****Statement of Informed Consent***

I agree to participate in a study about stress and coping during adolescence. The purpose of this study is to learn more about stressful life events, coping responses, and feelings of depression. I understand that the potential benefits of the study involve gathering information that can help to identify adaptive and maladaptive coping patterns, and thereby improve the ability to help people cope with various problems.

I agree to participate in this study for approximately 1 hour today. Participation in this study involves completing several questionnaires and a short interviews about depression and recent life problems. Also, my consent implies that the investigators can gather relevant information from my medical chart for use in the research study.

In addition to the information collected today, I will be contacted in 3 months for a brief reevaluation of my depression and the treatment I was provided. My consent now simply allows the investigators to contact me in 3 months. At that time, I can decide whether or not I want to continue my involvement in the study.

I am aware that the risks and consequences of my participation in the study primarily involve talking about emotional issues and protecting my confidentiality. Although it is unlikely that I will be exposed to any risks by participating in this study, you will ask a variety of questions about my life and my emotions. Some of the questions may force me to confront various emotions as I discuss these different issues.

All information collected in this study will be kept confidential. If the results of this study are published in a scientific journal, my identity will be kept anonymous and only group averages will be reported. No individual data or identifying information will be reported. Only the professionals involved in my treatment will be provided a summary of the results of my individual scores. The only other time the investigator would be obligated to disclose any information would be if I expressed specific plans for harming myself or another person or if reasonable suspicion of child abuse existed. If this happened, it would be necessary for the investigator to inform my primary physician or children's protection services in order to take any precautions deemed necessary.

I understand that participation in this study is voluntary. Even if I initially decide to participate, I can later change my mind and can choose to stop at any time. My decision of whether or not to participate will not affect my relationship with the hospital or its staff in any way.

My signature below indicates that I have read, understood, and agreed with the information contained in this consent form. Any questions I had were answered to my satisfaction. A copy of this consent form is available upon my request.

Patient _____ Date _____

Parent _____ Date _____

Investigator _____ Date _____

REFERENCES

- Abramovitch, R., Freedman, J. L., Henry, K. L., & VanBrunschot, M. (1995). Children's capacity to agree to psychological research: Knowledge of risks and benefits and voluntariness. *Ethics and Behavior*, 5, 25-48.
- American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, 47, 1597-1611.
- Appelbaum, P., & Roth, L. H. (1982). Competency to consent to research: A psychiatric overview. *Archives of General Psychiatry*, 39, 951-958.
- Brinkman, D., Overholser, J., & Klier, D. (1994). Emotional distress in adolescent psychiatric inpatients: Direct and indirect measures. *Journal of Personality Assessment*, 62, 472-484.
- Brinkman-Sull, D., Overholser, J., & Silverman, E. (1997). *Risk of future suicide attempts in adolescent psychiatric inpatients at 18-month follow-up*. Cleveland, OH: Case Western Reserve University.
- Brooks-Gunn, J., & Rotheram-Borus, M. (1994). Rights to privacy in research: Adolescents versus parents. *Ethics & Behavior*, 4, 109-121.
- Burbach, D. J., Farha, J. G., & Thorpe, J. S. (1986). Assessing depression in community samples of children using self-report inventories: Ethical considerations. *Journal of Abnormal Child Psychology*, 14, 579-589.
- Bussell, D. A. (1994). Ethical issues in observational family research. *Family Process*, 33, 361-376.
- Canter, M. B., Bennett, B. E., Jones, S. E., & Nagy, T. F. (1994). Ethics for psychologists: A commentary on the APA ethics code. Washington, DC: American Psychological Association.
- Ceci, S. J., Peters, D., & Plotkin, J. (1985). Human Subjects Review, personal values, and the regulation of social science research. *American Psychologist*, 40, 994-1002.
- Dekraai, M. B., & Sales, B. D. (1991). Liability in child therapy and research. *Journal of Consulting and Clinical Psychology*, 59, 853-860.
- Department of Health and Human Services. (1983, March 8). Protection of human subjects (Code of Federal Regulations, 45 CFR 46). Washington, DC: Author.
- Drotar, D. (1989). Psychological research in pediatric settings: Lessons from the field. *Journal of Pediatric Psychology*, 14, 63-74.
- Drotar, D., & Ievens, C. (1994). Preliminary report: Age differences in parent-child responsibilities for management of cystic fibrosis and insulin-dependent diabetes mellitus. *Journal of Developmental and Behavioral Pediatrics*, 15, 367-374.
- Drotar, D., & Robinson, J. (1998). Researching failure to thrive: Progress, problems, and recommendations. In D. Kessler & P. Dawson (Eds.) *Pediatric undernutrition and failure to thrive: A transdisciplinary approach* (pp. 74-92). Baltimore: Paul H. Brookes.
- Fisher, C. B. (1993). Integrating science and ethics in research with high-risk children and youth. *Social Policy Report. Society for Research in Clinical Development*, 7, 1-23.
- Fisher, C. B. (1994). Reporting and referring research participants: Ethical challenges for investigators studying children and youth. *Ethics and Behavior*, 4, 87-95.
- Fisher, C. B. (1998, August). *Discussion: Ethics code revision. Task Force discussion with ethical comments*. Presented at annual meeting of the American Psychological Association, San Francisco, CA.
- Freiheit, S., Overholser, J. C., & Brinkman-Sull, D. (1996). Utility of the Minnesota Multiphasic Personality Inventory personality disorder scales with adolescent psychiatric inpatients. *Journal of Youth and Adolescence*, 25, 803-815.
- Freiheit, S., Overholser, J. C., & Lehnert, K. (1998). Sense of humor and adolescent depression. *Journal of Adolescent Research*, 13, 32-48.
- Glantz, L. (1996). Conducting research with children: Legal and ethical issues. *Journal of the American Academy of Child and Adolescent Psychiatry*, 34, 1283-1291.
- Grodin, M. A., & Glantz, C. H. (1994) (Eds.). *Children as research subjects*. New York: Oxford University Press.
- Gray, B. H. (1977). The functions of Human Subjects Review Committees. *American Journal of Psychiatry*, 134, 907-910.
- Harth, S. C. & Thong, Y. H. (1995). Parental perceptions and attitudes about informed consent in clinical research concerning children. *Social Science and Medicine*, 40, 1573-1577.
- Hoagwood, K. (1994). The certificate of confidentiality at the National Institutes of Mental Health. Discretionary considerations and its applicability in research on child and adolescent mental disorders. *Ethics and Behavior*, 4, 123-131.

- Hoagwood, K., Jensen, P. S., & Fisher, C. B. (1996). *Ethical issues in mental health research with children and adolescents*. Mahwah, NJ: Laurence Erlbaum.
- Jonsen, A. R. (1978). Research involving children: Recommendations of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *Pediatrics*, 62, 131-136.
- Keith-Spiegel, P., & Koocher, G. P. (1990). *Children, ethics and the law*. Lincoln: University of Nebraska Press.
- Koocher, G. P., & Keith-Spiegel, P. (1998). *Ethics in psychology. Professional standards and cases*. New York: Oxford University Press.
- La Greca, A. M. (1990). *Through the eyes of the child. Obtaining self-reports from children and adolescents*. Needham Heights, MD: Allyn & Bacon.
- Liss, M. B. (1994). Child abuse: Is there a mandate for researchers to report? *Ethics and Behavior*, 4, 133-146.
- Lowe, C. U., Alexander, D., & Mishkin, B. (1974). Nontherapeutic research on children: An ethical dilemma. *Journal of Pediatrics*, 84, 468-472.
- Melton, G. B., Koocher, G. P., & Saks, M. J. (1983). *Children's competence to consent*. New York: Plenum Press.
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Morrow, G. R. (1980). How readable are subject consent forms? *Journal of the American Medical Association*, 244, 56-58.
- Nelson, S. H., & Grunebaum, H. (1972). Ethical issues in psychiatric follow-up studies. *American Journal of Psychiatry*, 128, 1358-1362.
- Novick, M. R. (1981). Federal Guidelines and Professional Standards. *American Psychologist*, 36, 1035-1046.
- Ogloff, J., & Otto, R. (1991). Are research participants truly informed? Readability of informed consent forms used in research. *Ethics and Behavior*, 1, 239-252.
- Overholser, J. C., Adams, D., Lehnert, K., & Brinkman, D. (1995a). Self-esteem deficits and suicidal tendencies among adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 34, 919-928.
- Overholser, J. C., Brinkman, D., Lehnert, K., & Ricciardi, A. M. (1995b). Children's Depression Rating Scale—Revised: Development of a short form. *Journal of Clinical Child Psychology*, 24, 443-452.
- Phillips, D. F. (1996). Institutional review boards under stress: Will they explode or change? *Journal of the American Medical Association*, 276, 1623-1626.
- Range, L. M., & Cotton, C. R. (1995). Reports of assent and remission in research with children: Illustrations and suggestions. *Ethics and Behavior*, 5, 1-13.
- Robinson, J. F. (1998). *Problem-solving abilities of mothers of infants with nonorganic failure to thrive*. Unpublished Ph.D. dissertation, Case Western Reserve University, Cleveland, OH.
- Roth, L. H., & Appelbaum, P. S. (1983). Obtaining informed consent for research with psychiatric patients. *Psychiatric Clinics of North America*, 6, 551-565.
- Rotheram-Borus, M. J. & Koopman, C. (1994). Protecting children's rights in AIDS research. In B. Stanley, & J. Sieber, (Eds.), *Social research on children and adolescents. Ethical Issues* (pp. 143-161). Newbury Park, CA: Sage.
- Ruccione, K., Krimer, R. T., & Moore, I. K. (1991). Informed consent for treatment of childhood cancer: Factors affecting parents' decision making. *Journal of Oncology Nursing*, 813, 112-121.
- Rutstein, D. D. (1969). The ethical design of human experiments. *Daedalus*, 98, 523-541.
- Scarr, S. (1994). Ethical problems in research on risky behavior and risky populations. *Ethics and Behavior*, 4, 147-155.
- Schwartz, C. E., Chesney, M. A., Irvine, M. J., & Keefe, F. J. (1997). *Psychosomatic Medicine*, 59, 362-371.
- Schwitzgebel, R. K. (1978). Suggestions for the uses of psychological devices in accord with legal and ethical standards. *Professional Psychology*, 9, 478-488.
- Scott-Jones, D. (1994). Ethical issues in reporting and referring in research with low-income minority children. *Ethics and Behavior*, 4(2), 97-108.
- Stanley, B., Sieber, J. E., & Melton, G. B. (1987). Empirical studies of ethical issues in research. A research agenda. *American Psychologist*, 42, 735-741.
- Stanton, A., Burker, E., & Kershaw, D. (1991). Effects of researcher follow-up of distressed subjects: Trade-off between validity and ethical responsibility? *Ethics and Behavior*, 1, 105-112.
- Steiner, I. D. (1972). The evils of research. *American Psychologist*, 27, 766-768.
- Stuart, R. B. (1978). Protection of the right to informed consent to participate in research. *Behavior Therapy*, 9, 73-82.

- Thompson, R. A. (1990). Vulnerability in research. A developmental perspective on research. *Child Development*, 61, 1–16.
- Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment*, 53, 827–831.
- Weithorn, C., & Campbell, S. B. (1982). The competency of children and adolescents to make informed treatment decisions. *Child Development*, 53, 1589–1598.
- Weithorn, L. & Schearer, D. G. (1994). Children as research subjects. In M. A. Groden & C. H. Glantz (Eds.), (1994). *Children as research subjects* (pp. 100–125). New York: Oxford University Press.

V

Disseminating Research Findings

One of the most important set of skills for researchers concerns the ability to disseminate information from research in presentations and written articles. Presentations are a vehicle to disseminate research findings to one's colleagues for critical review and constructive criticism. They also provide a way to obtain feedback on ongoing work, to apprise colleagues of one's research interests and progress, and to communicate the results to collaborators in medical and agency settings, which is important to develop research related collaborations.

The preparation of original research articles is especially critical to the professional development of scientist practitioners in pediatric and clinical child psychology. Researchers' success in academic settings depends on their abilities to make consistent contributions to the scientific and professional literature. Writing is also a primary vehicle that is used to communicate with colleagues about one's work and also provides a unique way to clarify one's thinking. As a complement to the contributions of original research reports, research reviews can be an important vehicle of defining and clarifying scientific problems, summarizing previous investigations to inform the readers of the state of the art in the area of science, and suggesting an agenda to give direction to the necessary ability developments to next steps of scientific inquiry.

Despite the importance of disseminating research findings, through presenting data at scientific meetings and by preparing original research articles and reviews, opportunities for students and practicing researchers to receive sufficient training to master these skills are relatively limited. To address this need, each of the chapters in Section V describes practical strategies to help readers disseminate research information from their studies. The preparation of these chapters offered Drotar an opportunity to summarize what he had learned from his own experience in presenting his research, preparing empirical journal articles, and developing research reviews. In Chapter 15, he describes the ingredients of an effective scientific presentation including the need to understand one's audience, developing effective take-home messages, putting information in slides and handouts, and presenting data.

Chapter 16 takes the reader through a step-by-step description of writing research articles for publication, including the process of manuscript review. It

describes the specific sections of a manuscript and considers what reviewers look for in a scientific manuscript. Specific examples of methods of presenting information in a research report and in issues responding to reviewers' in a revised manuscript are also presented.

Chapter 17 describes the special considerations that are necessary in writing research reviews, including a step-by-step approach for investigators in deciding on the focus of their reviews and how to structure their review article. The chapter also deals with special dilemmas in writing a review article such as articulating the novel contribution of the review, reviewing critically yet constructively recording information from studies that are reviewed, and so forth.

Chapter 18 will be very useful in helping readers to evaluate the quality of reviews, especially meta analyses. Durlak considers the role of hypotheses in reviews, representativeness of studies that are reviewed, appropriateness of analyses, issues in evaluating the practical significance of scientific data, and the meaning of outcomes that are presented in research that is summarized in meta-analysis.

In Chapter 19, final chapter in this section, Droter considers issues in reviewing and editing manuscripts for scientific journals. Detailed descriptions of editors' and reviewers' roles and responsibilities, the structure of journal editorial boards, and the characteristics of effective reviews are presented in detail. Ethical issues in reviewing and editing manuscripts such as managing conflicts of interest in the review process and ways of improving the review process also are considered.

Each chapter in Section V also considers recommendations for training researchers to develop the skills that are necessary to present their work in scientific meetings and to publish their work in scientific journals. In particular, the role of a writer's workshop seminar that involves ongoing group critique of scientific manuscripts, review articles, and presentations is described in detail as a useful way to help students develop their writing skills and identities as scientific writers.

15

Presenting Scientific Data

DENNIS DROTAR

Pediatric and clinical psychologists' presentations of scientific findings to their professional colleagues are important professional activities for several reasons: Presentations are a primary vehicle to disseminate research findings to one's peers for critical review and constructive criticism. While a clear, cogent presentation can do much to enhance colleagues' interest in one's research, a problematic presentation can detract from otherwise well-designed and well-executed science. Presentations also provide a way to obtain feedback on ongoing work, to apprise colleagues of one's research interests and progress, and to communicate the results of research to professional collaborators in medical and agency settings, which is important to sustain research-related collaborations (Drotar, 1995).

Despite the many good reasons for researchers to develop skills in presenting their data to colleagues, opportunities to obtain formal training in presenting scientific data are limited, both in graduate programs and continuing professional education. Consequently, most researchers learn to present their data through trial and error, by obtaining informal feedback from their audiences and colleagues, and/or by learning vicariously, that is, by observing their colleagues' presentations.

In order to develop and sharpen skills in presenting scientific data, there simply is no substitute for practical experience in presenting to a wide range of professional audiences. In addition, information about guidelines and helpful hints about making scientific presentations can be useful, not only to novice researchers who are developing their presentation skills but to researchers who want to refine their skills. The purpose of this chapter is to provide such guidance. Interested readers might wish to consult Aarons and Dunwoody (1995), Teel (1990), Thompson, Mitchell, Halvorsen, Foster, and Roberts (1987), or Williams (1995) for other relevant sources on this topic.

DENNIS DROTAR • Department of pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

THE TASKS INVOLVED IN MAKING A SCIENTIFIC PRESENTATION

Investigators who are interested in presenting their data at a scientific meeting face a number of different tasks and decisions, which are summarized in Table 1. These tasks include making decisions about whether to present data, where to present, and the various responsibilities of the presentation itself.

SHOULD MY DATA BE PRESENTED AT A MEETING?

Ideally, the decision to present data at a scientific meeting should be governed by an affirmative answer to the following basic, albeit potentially threatening question: "Are my data sufficiently interesting, important, and well executed to warrant presentation?" Nevertheless, in the real world, the decision about whether to present one's data can be governed by factors other than the readiness and quality of the information that is to be presented. For example, presentations of scientific data are ways to help establish and maintain highly valued social and professional contacts with colleagues. On a practical level, presentations also provide a way to lay claim to travel funds to support attendance at scientific meetings, which are increasingly hard to obtain in many work settings.

As tempting and as advantageous as it might seem to submit an abstract to a professional meeting, researchers should carefully consider several questions before deciding to present their data: (1) How developed will my data be at the time of the presentation? (2) Is there sufficient time to allow for adequate data analysis, consideration of the meaning and implications of the findings, and preparation of the presentation? (3) What other professional and/or personal activities overlap with the time of the meeting that compete for the time needed to prepare the presentation?

While the time involved in actually making a scientific presentation is short, the time that is necessary to prepare an effective presentation can be considerable, usually more than most investigators initially realize. Moreover, investigators need to make many difficult decisions concerning such questions as what information to include or not include in the presentation, preparation of slides, and so forth, which are not readily apparent at the time that one submits an abstract. Because it is easy to underestimate the time that is needed to adequately prepare to present data at a scientific meeting, investigators should leave themselves maximal preparation time.

WHERE SHOULD DATA BE PRESENTED?

Given ever-increasing limits on travel funds, combined with expanding options for professional meetings, researchers need to be increasingly selective about their conference attendance. What should one consider in making the decision concerning where to present data? Several factors are relevant, such as personal interest in the theme or topic of the meeting, prior participation in specific meetings, the audience one wants to reach with one's research, the relevance of the

Table 1. Checklist of Relevant Tasks for Scientific Presentations

Making decisions about the presentation
Whether to present
How developed will the data be for the presentation?
Is there sufficient time to prepare?
What other professional and personal activities overlap with the time of the meeting?
Where to present
Type of audience one wants to reach
Personal interest in the meeting
Relevance of research to conference topics
General responsibilities of presenters
Prompt attendance
Inform relevant person if unable to attend
Keep to time limits
Send copy of presentation/poster to those who request it
Find out essential details about presentation
Time
Place
Length of presentation
Size of room (if possible)
Audience (e.g., How many are expected to attend the presentation? What professional disciplines are represented? What is their level of training?)
General guidelines for presenters
Rehearse and time the talk
Prepare user-friendly slides and overheads
Carefully integrate text with slides
Stay within time limits for talk
Carefully field questions from your audience
Send papers to those who request them
Planning the focus of the presentation
Determine central question on theme of your presentation
Identify data that pertain to this theme
Decide what aspects of your data the audience will be most interested in
Focus presentation on these data
Planning the components of the presentation
Introduction
Need for the study
Purpose/research question
Hypotheses
Use tables to summarize points
Method
Design
Sampling/procedures
Description of sample
Measures
Use tables to summarize methods
Results
Summary of essential findings
Use figures and tables to highlight findings
Summary and implications
Summary of major findings
Clinical, theoretical, and research implications
Limitations
Directions for future research
Use tables to summarize points

research to the conference topics and themes, and the perceived importance of interacting with specific colleagues.

Researchers who work with pediatric and child clinical populations have a wide range of options concerning meetings and conferences to present their data. These range from the national American Psychological Association (APA) convention to specialized national conferences, such as the Florida Conference on Child Health Psychology or the Kansas Conference on Clinical Child Psychology. Pediatric and clinical child psychologists also have the option of presenting their data to interdisciplinary audiences, such as the Society for Research and Child Development, the Society for Behavioral and Developmental Pediatrics, the Society for Behavioral Medicine, or the American Association of Behavioral Therapy.

Regional conferences, such as the Great Lakes or Southwest Society of Pediatric Psychology meetings, provide another choice for pediatric psychologists, especially for researchers who prefer a smaller audience. Other options include specialized scientific meetings, such as the International Conference on Infant Studies or the Gatlinburg Conference on Mental Retardation. In my experience, specialized conferences have the important advantage of providing greater opportunities for interchange with colleagues, which can be especially useful to young researchers.

HOW ARE SCIENTIFIC PRESENTATIONS REVIEWED?

Presentations of data at scientific meetings fall into four basic categories: (1) poster presentations; (2) presentations of single papers, which are typically 10–20 minutes in length; (3) symposia in which papers with a common theme are presented; and (4) invited presentations, which are typically 30 minutes to an hour (Poling, Methot, & LeSage, 1985). The specific allocation of submissions to different formats (e.g., posters, presentations, or individual papers vs. symposia) is determined by a program or review committee, which is responsible to the sponsoring organization and is charged with reviewing and evaluating the submitted papers to determine which ones should be accepted and/or presented as papers versus as posters.

While there is some variation in the criteria for acceptance of scientific presentations, depending on the specific meeting, specific features of papers that are most important to reviewers include: (1) potential relevance of the research topic to conference participants interests, and/or theme of the conference; (2) the significance of the research and findings to the field; (3) adequacy of the research design; and (4) organization and clarity of writing (see Chapter 16, this volume).

Reviewers, who are typically blind as to the author's identity, are asked to assign scores for separate dimensions of their evaluation of abstracts that are submitted, such as significance of the research and organization and clarity of writing, as well as to provide an overall rating of quality. The number of abstracts assigned to individual reviewers varies widely as a function of a scientific organization, the number of submissions to the conference, and the availability of reviewers. Typically, reviewers evaluate from three to ten abstracts each, and more in some cases. For example, in an effort to enhance the consistency of the review

process across different raters, for several years each member of the program committee of the Society for Behavioral Pediatrics reviewed every abstract that was submitted to the meeting (typically 75–100).

Given the time constraints that are typically involved in planning scientific programs, reviewers are often under considerable pressure to review abstracts quickly, often within a week of receiving them. For this reason, they place a high premium on the clarity and reading ease of the abstracts as well as the interest that they believe the presentation will generate. Consequently, would-be presenters face significant challenges to make a clear and convincing case for their work in a brief format, which is typically between 100–300 words. Given the constraints of the format, careful attention to preparation of the abstract in accord with the specific guidelines that are distributed by the organizers of scientific meetings is critical (Poling et al., 1995). As is true of manuscripts, reasonably good scientific work may be turned down for presentation because it is poorly written or because the authors have not attended carefully to the specific guidelines for preparation of abstracts.

The mean numerical ratings of the abstracts assigned by the reviewers are used to make a decision concerning acceptance of a paper for a scientific meeting. The number of papers, symposia, and so on, that are accepted for a scientific program will vary as a function of the number of hours in the program that need to be filled compared with the number of submissions to that particular conference. Conferences that have a large number of slots for presentations will accept a majority of abstracts, especially if the author is willing to present a poster if work is not accepted as a paper. Nevertheless, in my experience, even if a generous number of potential presentation slots are available, very poorly organized or executed studies generally will not be accepted.

PRESENTERS' RESPONSIBILITIES AND TASKS

General Considerations

To help presenters understand the basic tasks that are involved in presenting scientific data, a summary of issues to consider is listed in Table 1. Once his or her paper is accepted for a meeting, the author's most obvious obligation is to show up and deliver the presentation. In the event that unforeseen circumstances prohibit one's attendance, it is the presenter's responsibility to arrange for a colleague to present the paper (or poster) in his or her place and inform the appropriate program representative about the change. One does occasionally encounter missing presenters who have not informed anyone about their absence. Such behavior not only is frustrating to the chair of the paper session or symposium but also to colleagues who may have arranged their time in order to hear the presenter.

Presenters also should carefully follow the guidelines for posters and paper presentations, for example, format, length, and style, that are provided by the conference organizers. Symposium participants should contact the organizer of the symposium prior to the meeting in order to obtain information concerning the time, allocation, and order of the presentations, if it is not clear from the program. Some

symposia organizers prefer to give an orientation to presenters and in some cases meet with the participants prior to their presentations. Such preparation and dialogue can be quite helpful, but participants' busy conference schedules often preclude such meetings. Discussants of symposia generally prefer to receive individual papers in advance of the meeting so that they have sufficient time to prepare their remarks. If at all possible, presenters should accommodate their discussants' preferences.

On the day of the presentation, presenters should arrive in sufficient time to arrange their slides (bringing a tray with slides already arranged will facilitate this task) and/or to ensure that someone is prepared to present their overheads and so forth. While close attention to the details can limit some problems with presentations, presenters cannot anticipate all contingencies. For example, this author's (Drotar) stint as a discussant at a regional pediatric psychology conference was once preempted by a tornado warning, which sent everyone scurrying from the conference room.

Preparing for One's Audience

The essence of an effective scientific presentation is to translate the purpose, method, findings, conclusions, and implications of a scientific data to an audience that is not necessarily knowledgeable about the specific topic of the research. For this reason, considerable thought and ingenuity are necessary to identify the most important, interesting features of one's research and translate them clearly to professionals who may not be familiar with the content or methods of the research that is to be presented.

To facilitate their primary task of tailoring their communications to a professional audience, presenters need to find out about their audience. One basic question is, how many people will be attending the presentation? Presenters should approach their presentations differently in a small group format compared with the larger audiences typically found at the APA meetings. Although it is usually difficult to anticipate the exact number of people that will be attending a specific presentation, presenters can sometimes gauge the size of the audience by finding out how many people are signed up for their session or seminar.

The audience's level of training will influence their interest, understanding, and needs, and hence should influence the presenter's approach. For example, in presenting research data that pertain to a clinical problem such as attention deficit with hyperactivity disorder, one might develop the content of the presentation differently for an audience of graduate students or residents who have seen relatively few children with this problem versus a group of experienced psychologists and pediatricians. Similarly, one would emphasize different aspects of methods and findings for an audience of undergraduate psychology majors versus a consensus conference of one's peers.

Unfortunately for presenters who are interested in tailoring their talk to their audience, audiences at scientific meetings tend to be heterogeneous in their experience, nature, and level of training and sometimes in professional discipline. In some cases, presenters cannot assume that the audience has a great deal of prior knowledge about their research topic. This means that presenters may need to give

a sufficient level of basic information about their research, for example, rationale, design, and basic findings, to allow those who are relatively unfamiliar with the work to follow the presentation. Although this strategy runs the risk of leaving some of the more experienced researchers in the audience somewhat disappointed, at least with certain aspects of the talk, the "basic information" approach is probably the most effective for the typical audience. The alternative strategy of assuming that one's audience is familiar with the research topic and methods and giving an in-depth, highly specialized talk runs the risk of losing one's audience or, worse yet, not engaging them in the first place.

Perhaps the most important characteristics of an audience that need to be considered are their professional training, affiliation, and experience, which can have a profound influence on their interests, needs, and understanding of information that is presented. In this regard, it is always easier to present to a professional audience who shares one's basic assumptions concerning the need for and importance of the topic, methods of measurement, data analysis, or implications of the findings.

In contrast, presentations (scientific or otherwise) to members of different professions are generally much more difficult because one's basic assumptions about the topic, concepts, or language may not be shared. For example, in giving presentations to pediatricians over a number of years, this author has learned through experience that this professional group is most interested in the practical and clinical implications of whatever data that are presented. Many pediatricians and other clinical practitioners are remarkably unpersuaded by what psychologists may regard as elegant statistical analyses or theoretical constructs. Consequently, unwaried presenters to such audiences can be brought down from the ivory towers of their data by predictable but difficult questions such as, "how is this finding going to help my patient" or, "how can I actually use this information in my practice?"

I recognize that one cannot always address such clinical questions with available scientific data. Moreover, in some cases it may not be responsible to do so. Nevertheless, anyone who presents to a pediatric audience or an audience of practitioners of any discipline should be prepared to discuss the potential clinical implications of whatever research they present and/or field questions concerning clinical relevance. Judging from the relatively limited emphasis on clinical implications of published research findings with pediatric populations (Roberts, 1992), the task of drawing out the clinical implications of research is an important one for any researcher who works with children and families, not only those who are presenting their data to clinical practitioners.

One salient advantage to presenting to audiences who do not share one's basic professional assumptions or language is that it forces presenters to carefully translate the details of their study and to draw out relevant implications for that particular audience. Knowing and respecting the professional affiliations of one's audience will help ensure the best possible fit between the language that is used in the presentation and the needs of the audience. It is easy to lapse into using profession-specific language when one is presenting to psychologists who are familiar with them. However, one profession's bread-and-butter language may be confusing to others. For example, pediatricians who are generally unfamiliar with such terms may be confused and/or frustrated by use in presentations unless the

terms are given suitable explanation. Upon hearing the term "path analysis" in a presentation, the pediatric practitioners of the 1990s may expect to see a description of care paths for a clinical problem, rather than statistical analyses.

Planning the Content and Structure of a Scientific Presentation

Once they have a reasonably clear idea of the nature of their audience, presenters need to plan the content and key messages of their presentation and consider how they will deliver them. The key operative word is "plan." There may be a few gifted orators who can wax eloquently on a moment's notice and plan their talk based on a few scribbled notes on the plane ride en route to the meeting or while listening to other presentations. However, for most of us, the quality of our presentations and our personal comfort are directly related to the level of our planning, preparation and thought.

The first and most obvious question for presenters is "what do I say?" The content of the scientific presentation includes both general issues that presenters would like to impart about the nature of their research area as well as more specific details about their research methods, data, and conclusions. The process of preparing a presentation involves a difficult series of choices about specific content, about how much to say about each individual area of the study, and how best to get one's messages across. In making such decisions, presenters should try to anticipate what specific aspects of their study and data will be most interesting to their audience and focus their presentation around these points, rather than to try to cover their topics too broadly. Because most presentations are given in highly restricted time formats (10–15 minutes is typical), researchers need to focus their presentations on the most salient findings, rather than attempt a comprehensive presentation of their data that may be well-intended but may only serve to frustrate their audience.

THE COMPONENTS OF A SCIENTIFIC PRESENTATION

To help structure and guide their presentations, it may be helpful for researchers to develop a brief outline of what they intend to present. As shown in Table 1, the components of a presentation include introduction of the need for the study, the purpose and scientific questions, methods, findings, and implications (see also Poling et al., 1995, for a detailed outline).

Introduction: Describe the Need and Purpose of the Study

The purpose of the introduction is to engage one's audience by informing them of the significance and basic purpose of the study. Presenters face the important challenge of engaging their audience quickly and convincing them it is better to stay and listen rather than use other potential options such as reviewing one's program or visiting the nearest Starbucks. One way to engage the audience's attention is to inform them about the need for the study by discussing clinical examples or briefly describing one's personal interest in the topic. It is important to make sure that your audience not only understands the significance of the study but also shares your enthusiasm for the need to conduct it. To accomplish this task, one

can underscore the importance of their study by discussing the numbers of children that are affected (in the case of clinical populations), describing controversies that remain unanswered, or identifying problems that arise from gaps in scientific knowledge that the study will address. The key features of the study's significance should be summarized on a slide or overhead.

Clarify the Primary Study Questions and Hypotheses

Once you have informed your audience about the need for and significance of your study, it is a relatively easy transition to describe the goal and purpose of your work and the scientific question(s) that will be addressed. Audiences are engaged by a clear statement of purpose and study question. The primary goals and questions that the study addressed should be described clearly and simply in several sentences and ideally summarized on a slide or overhead.

Because many studies in pediatric and clinical child psychology are complex and address several questions, presenters in these fields may face a difficult challenge in summarizing their research aims and questions. In this author's experience, effective presentations are generally built around central themes or questions that are introduced early and woven through the results and discussion of implications. A brief statement of hypotheses can be helpful to clarify the basic study question to one's audience and help the audience anticipate the data that will be presented. Clear visual diagrams or models of expected relationships among the variables of primary interest can also help the audience follow the findings that are presented.

Describe Basic Features of the Method

Because a presenter cannot possibly describe all the relevant methodological details of their research, they must make careful choices about what aspects of the method to describe. An audience generally wants answers to basic methodological questions such as: Is this a cross-sectional or perspective study? How many groups are involved? Who are the subjects? What outcome measures were used? How were the data analyzed (see Table 1)? Slides or overheads can be used effectively to underscore key aspects of study design, participants, measures, and analyses.

Present the Most Salient Data

Because all but the most parsimonious of researchers usually has more than enough data to present, they need to make difficult choices to avoid data overkill, which can confuse or even bore the audience. For this reason, presentations of data should focus on key questions and/or unusual findings of particular interest. Moreover, to help their audience follow their presentation of data, presenters should refer back to their hypotheses and provide summary statements.

Presentation of data should exhilarate, rather than exhaust one's audience. Cohen's (1990) dictum "less is more" should be followed religiously by presenters who may have the urge to share vast quantities of their data with an unsuspecting audience. In the interest of time and preserving the audience's attention and goodwill, it is not a good idea to present a great many secondary analyses.

Once having decided what data to present, investigators need to decide how best to present them, for example, whether to utilize figures, tables, or both, and what specific data to include on tables. Investigators should ensure that their tables and figures are audience-friendly, for example, they can be read and understood readily and do not contain too much information. Well-organized tables and figures can help presenters to take their audience step by step through their findings, while information-saturated slides can frustrate the audience. Because it is not easy to anticipate how tables and/or figures will look on slides, presenters should plan sufficient time to allow them to review alternative formats of tables and figures before they settle on one that they believe will be the most audience-friendly.

Give a Cogent Summary of Conclusions and Implications

Effective presenters will help their audience understand the meaning and implications of their central findings. To accomplish this task, it is generally useful to summarize the major findings or conclusions on slides and highlight the basic, "take-home" implications for one's audience. Audiences also appreciate some discussion of implications of data that are presented for development of theory, new research, or for practice. Many experienced presenters also discuss qualifications or any limitations of their study that affect interpretation of the data. Such information helps one's audience place findings in a context and minimizes misinterpretation.

Finally, most audiences are generally interested in hearing about the new scientific questions that are generated by research findings. One of the primary benefits of presentations is that they stimulate colleagues' ideas for future research or new applications of methods. For this reason, presenters should make sure that their audience receives the benefit of their thinking about the next steps of the research and future directions for the field.

GENERAL GUIDELINES TO ENHANCE THE CLARITY AND VALUE OF PRESENTATIONS

This author's experience in preparing, conducting, and listening to scientific presentations has yielded some general guidelines that may be useful to others in enhancing the clarity and value of presentations to one's audience. These are summarized in the next section.

Rehearse and Time Your Talk

Presenters should limit their presentations to the time allotted. As a rule, brevity is much appreciated by the audience and fellow presenters, while long-windedness is not. Overly long presentations stem from the fact that presenters often have a great deal of information that they want to share with their audience but have not rehearsed the content and timing of their talk. Most presenters underestimate the amount of time it takes them to give a talk. For example, assuming an average speaker who is presenting at an average rate of 120 words per minute (assuming a reasonable, nonpressured delivery), a 15-minute paper will

take about 1800 words or about 7 pages of text. Given such constraints, presenters need to be alert to the fact that unless they happen to be unusually terse individuals, they will inevitably have to do some editing of their talk. In order to determine exactly how much editing is needed, there is no substitute for rehearsing and timing the presentation.

In addition to helping to obtain a realistic estimate of the timing of the talk, a rehearsal accomplishes several important goals including: (1) facilitating editing and decision making about the number of slides; (2) making decisions about points of emphasis; (3) preparing audience-friendly slides or overheads; and (4) helping the presenter to become as familiar as possible with the material, thus reducing anxiety.

Depending on the presenter's preference, a rehearsal can vary from the "privacy of one's office," in front of the mirror approach, to a formal presentation rehearsal with a relatively large audience and critique. Such a dress rehearsal is closest to the actual conditions of the presentation and probably involves the best learning experience and opportunity for feedback. Having recognized the advantages of such rehearsals, some departments of pediatrics insist on a formal rehearsal and critique of faculty presentations for scientific presentations at meetings such as the Society of Pediatric Research. However, because it may not always be feasible to convene such a group, some presenters may prefer a compromise plan of rehearsing their talk in front of a few trusted colleagues.

Given the press of competing professional and personal activities that are faced by most researchers and the power of procrastination, some presenters handicap themselves by not finishing the text of slides of the presentation in sufficient time to do any rehearsal, let alone use feedback from the rehearsal to modify their presentation. Unfortunately, such an approach may sometimes translate into a "let the audience beware" strategy.

Prepare User-Friendly Slides—Overheads

Careful presentation of visual aids such as slides or overheads is a cornerstone of an effective scientific presentation. In the absence of such visual guides, it is difficult if not impossible for an audience to understand the details of scientific data from an oral presentation. Advances in computerized slide technology now provide presenters with many options for preparing visually interesting slides, especially compared with previous technology. (Some of us remember the advent of the blue background as an exciting advance over the black and white slide!) However, while modern slide-making technology is extremely helpful, it does not change the presenter's basic tasks, which are to decide on the most effective method of imparting data or information concerning methods and data in verbal and visual form, organize the presentation into clear points of information, place information into an effective sequence, and coordinate the text presentation with the slides.

Choosing Type and Format

One of the first tasks that presenters face is the choice of visual aids (e.g., slides, overheads, handouts, videotape, or multimedia). Individual presenters may have different preferences for various formats. For example, as a psychologist who has

spent most of his professional career working in a medical setting, this author prefers slides and rarely uses overheads. However, many psychologists in academic departments of psychology prefer overheads, which are easier to make and less expensive. The specific format of visual aids, for example, slides or overheads, is probably not as important as the clarity of the information that is presented.

Presenters should determine which type of format for visual aids works best for their purposes and audience. When they are prepared effectively, slides have the advantage of an engaging visual presentation as well as automated presentation. On the other hand, experienced presenters who have suffered through technical disasters with slides may not be so enchanted with these advantages. Moreover, presenters should be careful not to get too carried away by the potential of computerized slide technology. It is now possible to develop slides with such exotic color combinations and striking graphic displays that they actually distract the audience's attention away from the content of the presentation.

Selecting Content

The purpose of using slides or overheads is to clearly highlight points that the investigator wishes to emphasize to the audience. This can be done by using text selectively and by summarizing data in tables and figures. Because information is used to highlight points to one's audience, all slides or overheads should be self-explanatory and have a clear message.

By far, the most common problem encountered in slide-based presentations involves putting so much information on a slide that it cannot be easily understood by the audience and in some cases cannot even read. This problem unfortunately has become part of the culture of scientific presentations. For example, how many of us can recall sitting through presentations that included impossible to follow slides of a 10×10 correlation matrix or the "grand table" of 20 means and standard deviations? Speakers will sometimes try to cushion the visual distress and boredom of the jam-packed slides with cautionary statements such as, "You probably are not going to be able to read this," or the more understated apology, "I know that this slide is a little busy, sorry."

The prevention of information overload on slides or overheads is clearly a much better strategy than the use of apology. Unlike other problems that may plague presenters, such as inconclusive findings, the problems that are posed by information-saturated slides are preventable by rehearsal and consultation with peers. To be fair, one cannot always anticipate the size of the room, which can contribute to the problem of slide legibility. Nevertheless, presenters always can anticipate this problem by putting less information on individual slides. Those who want to give their audiences more detailed information concerning their sample, method, or findings can exercise the option of preparing handouts.

Integrating Slides and Prepared Text

One important challenge for presenters is to coordinate the information on slides or overheads with the text of their presentation. Presenters have several different options for preparing their text and integrating it with the slides or overheads they have prepared. The completely slide-based talk has the potential advantage of helping to maintain audience interest. On the other hand, one dis-

advantage of this approach is the difficulty of managing the presentation without the guidance of text. For example, the presentation can become completely focused on the slides. Some presenters are tempted to read from their slides, which is not interesting or engaging. Finally, the slide-based talk requires the presenter to possess a high level of familiarity with the content of his or her slides, which is not always the case, depending on how many times the presenter has given the talk.

Many presenters use a combination of prepared text and slides, which, in the right hands, can make for an effective presentation. In order to manage this type of presentation, presenters need to be thoroughly familiar with the text. Otherwise, they will be forced into reading a “speech,” which is always less interesting to the audience. Moreover, when using prepared text, presenters need to ensure that the sequence of their text fits their slides. Consequently, the presenter’s text should clearly indicate when the next slide should be shown. I also have found it helpful to decide what specific points will be made about each of the slides and put them directly into the text.

Field Audience Questions Carefully and Respectfully

Fielding questions from the audience can be a difficult, intimidating experience, especially for novice presenters. Presenters understandably dread questions that may expose flaws in their work or question assumptions of their work. Nevertheless, rather than anticipate the worst-case scenarios from audience questions, it is much more helpful for presenters to interpret questions as indications of interest, which in fact, is often true. Moreover, questions provide an opportunity for learning from one’s audience and having a dialogue about one’s work, which is very much in the spirit of scientific inquiry.

Two guidelines that this author has found to be useful in managing questions are: (1) understand the question before answering it; and (2) try not to be defensive in responding. In the heat of an anxious moment at a presentation, it may not be easy to understand what is meant by the question. Consequently, it is often useful to ask for clarification of a question that may be confusing. Answering an audience’s questions in a straightforward, nondefensive manner is always a good policy, albeit one that is not always easy to implement. Presenters should remember that it is always permissible to respond to a difficult question with such comments as: “that’s a good question; I really don’t know the answer, but I do have a few thoughts about it.”

Send Papers to Those Who Request Them

One of the professional courtesies, if not responsibilities, that presenters have is to send copies of their paper to those audience members who request them. Such information can be especially important for psychologists who work in relative isolation, for example, in practice or at small universities as a source of ideas and methods (Rienzi and Allen, 1994). While sending a copy of a poster or presentation would appear to be a straightforward professional responsibility, apparently it is not the norm for psychologist presenters. Rienzi & Allen (1994) found on average a response rate of only 51% for poster presentations across three different professional meetings of psychologists. In some cases, there may be legitimate reasons for researchers not to send papers based on their presentations, for example, the text

was not written out and was based entirely on slides. Nevertheless, in the interests of professional courtesy, it is often possible for researchers to send related material, copies of the slides, or a letter explaining the circumstances to interested parties.

COMMON PITFALLS IN GIVING PRESENTATIONS

One difficulty for presenters is that specific feedback from one's audience is not always apparent. With the exception of obvious cues, for example, large numbers of people who are in various stages of sleep (which may not be detectable in the darkened atmosphere of a slide presentation) or a mass exodus of their audience, presenters often do not have effective ways to gauge how their presentation was received. However, in accord with requirements for continuing education, conference organizers solicit specific feedback about the quality of presentations. While such feedback can be sobering for presenters to read, it can be very helpful to speakers who wish to improve the style, process, and/or content of their presentations. Unfortunately while such specific feedback is not consistently communicated back to presenters, it can be obtained from the conference organizers.

As one example of a model for such feedback, the Society for Behavioral and Developmental Pediatrics gathers detailed audience feedback on individual presentations at their annual meeting, summarizes the information, and feeds it back to presenters. Meeting participants are asked to provide information concerning content, delivery, and slides. The following is a summary of typical comments made about individual presentations from a recent meeting in each of these areas.

Problems in the Clarity and Relevance of Content

Several comments had to do with problems in the clarity of the content of the presentation, for example, "it would have been extremely helpful if the potential clinical relevance of the study had been stated at the outset, it was extremely hard to follow the presentation without this information"; or (my favorite): "I didn't have a clue what he was talking about." Other conference attendees commented on the speaker's failure to clarify the purpose of the research, for example, "what is the conceptual basis of the study"; problems with organization: for example, "it was not clear what the two parts of the talk had to do with one another"; and other specific problems with clarity, such as: "the study question was not clear; patient populations should be defined more clearly; terms were not clear; too much statistical jargon."

Problematic Delivery

Descriptions of presenters "going too fast" was the most common single complaint about their deliveries. Other general concerns about the delivery of presentations included, "mumbled, confused delivery," "never looked up," "interesting material but was not apparent in tone of the presenter," or "speaker seemed nervous, should have practiced." Another set of delivery-rated comments had to do with the presenter's management of slides, for example, "do not apologize for slide," "need to discuss slides," "don't read slides literally," "too fast over the slides," "too many slides," or "need more slides."

Inconsistent Quality of Slides

A final set of comments referred to problems with specific details of slides, for example, “the speaker used identifying labels that are not generally known, making the presentation hard to follow;” “colored slides were too hard to read;” “slides not projected well;” “star in upper left corner of slides was distracting;” “too much on slides;” and “abbreviations on slides were not clear.”

POSTER PRESENTATIONS

Posters have become a popular method of presenting scientific information at scientific conferences. In fact, posters are the most frequently used method of presenting information because a large number can be given at the same time. For this reason, many researchers’ initial experiences in presenting their scientific work to their peers will involve posters. Posters also have a clear advantage for researchers who dislike public speaking and/or welcome the opportunity to discuss their research with interested colleagues at a more leisurely pace than is possible during presentations.

General Information about Poster Presentations

In most scientific meetings, posters are typically grouped around a broad theme, for example, chronic illness in children, which can still include a heterogeneous array of presentations. Posters are typically displayed for 1–2 hours in a large meeting room. Researchers who prepare posters are required to prepare the information for the poster that describes their research methods and findings, mount it approximately in accord with guidelines, and stand by their poster to be available to discuss the information with colleagues who express interest in their work. Several guidelines (Poling et al., 1995) for poster presentations that are important to follow: (1) make the poster eye-catching; (2) make the poster ready to read; (3) cut detail to a minimum; (4) have details available for readers who desire them; and (5) have copies of a written abstract available and for manuscript with more details about the study available to distribute.

Organizing and Presenting the Poster

The information that is presented on a poster is organized in the following sections: title, abstract, introduction, methods, results, and discussion. This information follows the general format of a journal article but in highly abbreviated form. As is true for presentations, it is important to ensure that the poster is effectively organized so that the individual sections, introduction, methods, and so forth, logically follow one another. The challenge of preparing a poster is to boil the essence of the purpose, method, results, and implications of a study down into a few words. Because this task is not easy to accomplish, it is often helpful to work with a colleague who can critique the narrative and ensure that it is clear. To facilitate readability, information contained in posters can be presented in outline form. For example, study aims and implications can be enumerated or bullets can be used. Easy to read figures and/or tables should be used to summarize the sample and major findings.

As with presentations, a researcher's zeal to showcase a great many details of their results can overwhelm the unsuspecting poster reader, many of whom may already be bleary-eyed from their wanderings through the poster hall. Presenters of posters also should be prepared with additional information about the study, for example, an abstract and references, which conference attendees appreciate.

It is a considerable challenge to distill the essence of one's work into a few short paragraphs, tables, and figures for a poster; consequently, more time is required than most presenters anticipate to prepare an informative poster, especially if one has not had much experience with this format. Consequently, poster presenters should leave themselves ample time to prepare the text for their poster, prepare and consider alternative formats for figures and tables, and decide on the most appealing way to present their information. Specific instructions for preparation and layout of posters generally will be given by the conference organizers, and these should be followed. Conventions generally provide large poster boards that mount on easels. Presenters usually bring their posters in smaller pieces and attach them to the poster boards with pins. Because they may not be available, it is usually a good idea to bring pins for mounting the poster.

Authors who also are able to manage large quantities of paper can offer longer summaries and even manuscript versions of posters. However, presenters should be aware that at the large meetings even generous quantities of these documents will be quickly exhausted. Consequently, it is useful to have a sheet of paper handy so that colleagues can write down their names to obtain a summary. Sophisticated conference attendees will bring a supply of address stickers.

Another feature of presentation etiquette includes arriving on time to put up the poster and taking the poster down in a timely fashion. Because there is generally a small window of time—10 minutes or so—in between poster sessions, promptness and speed are of the essence in the world of the poster presenter.

Choosing the Style of the Poster

Poster presenters have their choice of a bewildering array of styles, ranging from the ultraminiinalist style (e.g., one occasionally sees typed manuscript pages attached to the board masquerading as a poster) to the high-tech, computer-generated, "seamless" poster. While poster presenters should settle on a style that fits their preference, esthetic sense, research, and budget, the minimalist approach should be rejected. Generally, a straight-forward, basic style poster is perfectly adequate so long as the information is clearly presented and can be read at a reasonable distance. In order to attract attention in a large poster session, it is particularly important that the title of the poster stand out at a distance.

TRAINING TO PRESENT AT SCIENTIFIC MEETINGS

Despite the importance of scientific presentations of data to enhance professional visibility and reputation, formal training in scientific presentations is limited or nonexistent in most graduate training programs and is difficult to come by for faculty who may want to polish their presentation skills. For this reason, there clearly is a need to develop such training, not only for students, but for faculty.

We have developed some training methods to help graduate students prepare for presentations at scientific meetings. Our students were interested in obtaining some training in preparing presentations in response to the opportunity and anxiety at having their papers selected to be presented at scientific meetings. As an introduction to this training, students were offered a lecture that summarized the salient points in giving presentations, many of which have been summarized in this chapter. The second and most essential part of the training experience involved supervised practice in giving presentations of their data in group format. The class was structured as follows: Each student was given 15 minutes to present information on their study using relevant handouts, slides, or overheads to a small-group audience of fellow students and a faculty mentor. This was followed by a 5-minute period where students responded to questions from the assembled group. Students who did not have a paper to present were asked to present their research proposals or other work in progress in the format of a presentation. The final part of the exercise was a group critique in which students and the instructor provided feedback concerning what they liked about the presentation and what they found problematic, and gave concrete suggestions, for example, content, delivery, and preparation of visual aids, to improve the presentation.

This training experience turned out to be a useful one for several reasons: Students were given practical experience in preparing and giving a talk in a setting where they experienced the support and feedback from colleagues. Another positive benefit of this exercise was that it provided a desensitization experience for students concerning the understandable anxieties that surround public speaking. Finally, the experience proved to be a useful vehicle to impart general guidelines concerning scientific presentations, as well as preparation of slides and handouts. Psychologists and mentors in professional settings can provide much the same support for their colleagues using a similar structure. We encourage would-be presenters to seek out such support.

REFERENCES

- Aarons, J. M., & Dunwoody, R. (1995). Developing a more viewable visual presentation. In M. W. Pequegnat & E. Stover (Eds.), *How to write a successful research grant application: A guide for social and behavioral scientists* (pp 177–183). New York: Plenum Press.
- Cohen, J. (1990). Things I have learned so far. *American Psychologist*, 45, 1304–1312.
- Drotar, D. (1995). Consulting with pediatricians. *Psychological perspectives*. New York: Plenum Press.
- Poling, A., Methot, L. L., & LeSage, M. G. (1995). *Fundamentals of behavior analysis research*. New York: Plenum Press.
- Rienzi, B. M., & Allen, M. J. (1994). Poster presenters. Send us your papers. *American Psychologist*, 49, 681–682.
- Roberts, M. C. (1992) Vale dictum: an editor's view of the field of pediatric psychology and its journal. *Journal of Pediatric Psychology*, 17, 785–805.
- Teel, C. S. (1990). Completing the research process: Presentations and publications. *Journal of Neuroscience Nursing*, 22, 125–127.
- Thompson, W. M., Mitchell, R. L., Halvorsen, R. A., Foster, W. L. Jr., & Roberts, L. (1987). *Investigatory Radiology*, 22, 244–245.
- Williams, J. B. W. (1995). How to deliver a sensational scientific talk. In W. Pequegnat & E. Stover (Eds.), *How to write a successful research grant application: A guide for social and behavioral scientists* (pp 171–176). New York: Plenum Press.

16

Writing Research Articles for Publication

DENNIS DROTAR

Writing is critical to the professional development of scientist-practitioners in pediatric and clinical child psychology in several respects: Professional success in academic settings depends on consistent contributions to the scientific literature. Writing is also the primary vehicle that is used to communicate with colleagues about one's research. Over and beyond the pragmatic importance of scientific writing to the professional careers of pediatric and clinical child psychologists, this endeavor provides an unique way to clarify one's thinking and is an intellectually exciting activity, which can provide a welcome alternative to the high-stress, action-oriented world of the practitioner.

Despite the importance of developing and sustaining productivity in writing for professional development, the productivity of clinical psychologists as assessed by the number of publications in referred journals has been less than optimal (Haynes, Lemsky, & Sexton-Radek, 1987). Norcross, Karg, and Prochaska (1997) reported that the modal number of articles for clinical psychologists who were surveyed was 0, with a median of 0.5 and mean of 1.9. To my knowledge, data concerning the publication records of clinical child psychologists have not been published. A survey of pediatric psychologists found that they had an average of 1.8 publications in the past year and 6.7 in the proceeding 5 years (Drotar, Sturm, Eckerle, & White, 1993). In evaluating these findings, readers should consider that the respondents were a relatively established group of pediatric psychologists who had an average of 10 years of post-PhD experience.

Constraints on the scholarly productivity of scientist-practitioners are complex and relate to a myriad of factors, especially competing work-related activities (Haynes et al., 1987). My own experience has revealed another significant con-

DENNIS DROTAR • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

straint on writing: Professional psychologists who work with children receive insufficient training in writing research articles either in graduate or postgraduate education. Graduate training concerning writing is focused primarily on writing papers for courses, master's theses, and dissertations. While these are important experiences, they are quite different written products and require somewhat different skills than manuscripts for journal articles (Bem, 1987). Consequently, many pediatric and clinical child psychologists may leave their graduate training without having sufficient familiarity with the process of submitting research-based articles to professional journals or sufficient experience and skills with writing such articles.

The purpose of this chapter is to describe lessons I have learned as a contributor to the research and professional literature in clinical child and pediatric psychology. Readers who are interested in cogent general discussions concerning scientific writing might wish to consult Becker (1986), Day (1988), Henson (1995), Howard & Barton (1986), Kellogg (1994), or Sternberg (1988). Sources concerning specific issues in publishing research reports include Home (1988), Maher (1978), Northridge & Susser (1994), Osipow (1996), Parry (1989), Peterson (1996), Roberts, Lyman, Breiner, and Royal (1982), Squires (1989, 1990), and of course the publications manual (American Psychological Association, 1994). Finally, readers also should read the companion chapter in this volume (Chapter 19) that deals with reviewing and editing manuscripts for scientific journals.

MOTIVATION FOR WRITING RESEARCH ARTICLES

Writing is an inherently difficult and time-consuming task for most of us. Moreover, exposing one's writing to the scrutiny of others is by no means devoid of stressors, although with experience one can learn to master them. What, then, are the positive benefits of writing that balance the disincentives associated with writing and can enhance one's motivation? Speaking from my own experience, a primary source of gratification for writing is the fact that it helps to structure and clarify my thinking (Henson, 1995). Simply put, I do my best thinking while writing.

Writing also provides a way to contribute to one's field in ways that are otherwise not possible through clinical work alone. Again, in looking back on my own career, one of the exciting advantages of writing has been the opportunity it has afforded to help develop and shape the emerging field of pediatric psychology through descriptions of practice and research (Drotar, Lemanek, LaGreca, & Kazak, 1995).

For those pediatric and clinical child psychologists who work in academic settings, motivation for writing includes direct, tangible rewards as well as considerable pressures. Productivity in publishing articles in scientific journals is a primary criterion for tenure and promotion at all universities. Moreover, in some settings, achieving promotion and/or tenure is tied directly to continuing one's position, to salary advancement, and, hence to professional satisfaction (Holden & Black, 1996). While one can debate the merits and drawbacks of the "publish or perish" dictum, the importance of publication for professional advancement in academic settings is not going to change any time soon, and hence is a powerful force to which psychologists in these settings must attend.

WHAT TYPE OF ARTICLE SHOULD ONE WRITE?

Although this chapter will be devoted primarily to the empirically based journal article, prospective writers should appreciate that they have a wide choice of formats, depending on what they have to say and the audience whom are trying to reach. The primary focus of the empirically based journal article is the presentation and synthesis of new scientific data. Other writing options include review articles, which provide a summary and synthesis of available research, new insights on a research problem, development of a new theory and/or new directions for research (see Chapter 17, this volume). Other options include case studies and series, which can provide useful descriptions of clinical phenomena and/or promising method of interventions (Drotar et al., 1995; see also Chapter 20, this volume). Finally, professionally oriented descriptive articles contribute a new slant on an ethical, training, or practice issues. See *American Psychologist or Professional Psychology, Research, and Practice* for examples of such articles.

CONSIDERATIONS IN PLANNING AN EMPIRICAL ARTICLE

Choosing the Focus

In his interesting, articulate account of preparing the empirical journal article, Bem (1987) noted that there are two possible articles to write: (1) the article that the investigator planned to write when the study was designed, and (2) the article that makes the most sense once the data are carefully reviewed and explored. Successful scientists let their data be the guide to their writing and focus their publications on the most interesting or important new findings that extend scientific knowledge in their field of research.

One of the most difficult issues in presenting data, especially from complex data sets, is determining the focus of the manuscript. Even experienced writers can be distracted by the options in presenting findings and alternative possibilities for the focus of an article. For this reason, it is important that a prospective author carefully select a major theme or focus for data presentation and tailor the manuscript around it. I recognize that in some cases it may be difficult for authors to determine what is in fact the most important or relevant aspect of their data for a particular audience. For this reason, consulting with colleagues and/or coauthors can be an invaluable way to help determine the most useful focus of one's manuscript.

For What Audience Should One Write?

Because each manuscript should be tailored to the target audience and specific goals of a particular journal, identifying one's potential audience and publication outlet are important for authors. Some manuscripts are rejected out-of-hand because the research does not fit directly with the journal's mission statement. Consequently, it is useful to have an audience in mind from the moment one starts to prepare a manuscript.

But how does one decide on a publication outlet for one's work? Several factors can influence an author's decision concerning a specific publication outlet, for example, the specific audience one wants to reach concerning their findings, peer

recognition, and/or professional advancement (Roberts et al., 1982). Some authors have a group of researchers in mind that they feel would be most interested in their specific findings. Others publish in journals they read regularly and/or where their colleagues or mentors publish.

One's professional setting also can make a difference in selection of a publication outlet. For example, psychologists who work in a medical settings often find it advantageous from the standpoint of collaboration with colleagues to publish in leading medical or interdisciplinary journals rather than in journals that are read mostly by psychologists. On the other hand, clinical child and pediatric psychologists in university psychology departments need to meet the demands of promotion and tenure requirements in their settings to publish in leading American Psychological Association (APA) journals. Given the high rejection rates of many journals, authors may wish to consider several options for publication rather than have their hearts absolutely set on one audience or outlet. Having several publication options will give authors more flexibility in deciding whether they should attempt a revision or choose another publication outlet in response to editorial feedback (Roberts et al., 1982).

In order to determine the degree of fit of their research to the journal's mission, prospective authors should carefully read the statements of journal editorial policies. Moreover, authors also will generally find it useful to consult recent issues of a journal of interest to obtain specific information about what kinds of articles are generally published. Having a thorough working knowledge of the kinds of articles that are published in a journal will help prospective authors make an informed decision about where to send their articles.

Journal Options in Pediatric and Clinical Child Psychology

So how does one locate a journal in which to disseminate one's research among the many possibilities (see above)? The fifth edition of *Journals in Psychology* (American Psychological Association, 1997) is a comprehensive resource that contains information concerning publishers, editors, and editorial policies for 355 journals. Authors also might consult Roberts and co-workers' (1982) description of publishing child-oriented articles in psychology, which contains a compendium of publication outlets, including various issues (e.g., publisher, coverage area, recent topics) for 106 journals.

Based on my experience, I also offer the following information concerning journals that publish research articles relevant to child clinical and pediatric psychology. For example, the *Journal of Pediatric Psychology* is a useful outlet for articles that focus specifically on research with pediatric populations, including assessment and intervention. The companion journal for research articles that focus on clinical child populations is the *Journal of Clinical Child Psychology*, which publishes original research, review and research articles, as well as work concerning training, advocacy, and professional practice in clinical child psychology.

The *Journal of Abnormal Child Psychology* publishes research concerning psychopathology in childhood and adolescence, especially empirical investigations in etiology, assessment, treatment in the community and correctional institutions, prognosis and follow-up, epidemiology, remediation in the education set-

ting, pharmacological intervention, and studies related to the ecology of abnormal behaviors.

Prospective authors whose research focuses on services related to children, policy trends, or program evaluation might wish to consider *Children's Services: Social Policy Research and Practice*, which is a new journal that will publish information concerning social policy issues related to services for children and families, particularly as they bear on mental health, psychological development, and welfare. Researchers who work in areas of child psychopathology or health who are interested in reaching a broader audience of psychologists with their work might want to consider such journals as *Journal of Consulting and Clinical Psychology*, *Journal of Abnormal Psychology*, *Behavioral Therapy*, *Child Behavioral Therapy*, or *Health Psychology*.

Authors who want their research with children to reach an interdisciplinary audience have several options. Those who want to reach pediatricians can submit their research to *Pediatrics* or the *Journal of Developmental and Behavioral Pediatrics*, which reaches a subgroup of pediatricians who have specialized clinical and research interests in behavior and development. Those who want to reach a very broad audience of professionals, for example, nurses, child life workers, social workers, and so on, who are interested in issues concerning the psychological care and health of children in health settings may wish to publish in *Children's Health Care*. Researchers whose work focuses on psychopathology and who want to reach a broader professional audience that includes child psychiatrists and social workers may want to consider the *American Journal of Orthopsychiatry*, *Journal of the American Academy of Child and Adolescent Psychiatry*, *Journal of Child Psychology and Psychiatry*, or *Clinical Child Psychology and Psychiatry*, a new multidisciplinary journal that is focused on clinical practice issues.

HOW ARE ARTICLES REVIEWED?

To be most effective in their writing, authors need to understand what is involved in the review process. Although there are some minor differences in individual journals' policies and practices, for the most part they all follow similar procedure.

One person, either the journal's editor or associate editor, or in some cases an editor who is appointed to manage a special issue determines the relevance of the submission to the journal's mission, assigns reviewers to provide critiques of the manuscript, makes the editorial decision, and writes the disposition letter to the author that details the editorial decision and suggestions for revision, if applicable. Depending on the journal's editorial policy, anywhere from two to five reviewers may review a manuscript.

What Feedback Are Authors Given Concerning Their Manuscripts?

Reviewers are responsible for contributing objective, balanced reviews of manuscripts in a timely (3–8 weeks, depending on journal policy) and ethical manner (e.g., impartial, confidential reviews). See Chapter 19, this volume, for more detail on reviewers' responsibilities.

Although specific evaluation forms differ from journal to journal, reviewers generally provide ratings of the quality of the manuscript in several key dimension as well as a narrative that summarizes their critique of the strengths and weaknesses of the manuscript. For example, the reviewer's recommendation form for *Journal of Pediatric Psychology* (JPP) (1997) includes the following categories: significance of the issue that the study addresses, rigor of method, logic of the author's reasoning, new knowledge contributed, clinical relevance, and other (to be specified by reviewer). Each category is rated on a scale of 0 = most negative to 9 = most positive. There also is a section for comments to the editor. In addition to making these ratings, reviewers for JPP are asked to indicate a recommendation for disposition from among the following choices: (1) acceptance unconditionally with only minor changes; (2) acceptance, either in present form or with minor revisions; (3) revise and resubmit in which there is merit to the manuscript, but sufficient level of revisions are needed to require another review; or (4) rejection, either with some reluctance or unqualifiedly. Moreover, if they recommend rejection of the manuscript, reviewers are asked to indicate the reasons that led to their decision, for example, insufficiently important, inappropriate for JPP, faulty in conception or design, poor presentation or writing, premature readiness for publication, faulty conclusion, and an "other" category. Reviewers also contribute a detailed narrative that describes their evaluation of the manuscript and suggestions to the author to improve it. Based on the above information and their own reading of the manuscript, the managing editor makes a decision concerning the manuscript's disposition, which is documented in a letter to the author.

Editorial Decisions

Three kinds of editorial decisions are generally rendered: accept with revisions, revise and resubmit, and rejection. While it is in principle possible for authors to receive an "accept with no revisions" editorial decision, in more than 24 years of submitting and reviewing manuscripts, I have not had direct experience with this particular option.

Based on my experience, manuscripts that fall into the "revise and resubmit" category include the following types: (1) those that have scientific merit based on their design and method, and so on, but contain writing that is so unclear that the essential scientific contribution is difficult to decipher; (2) manuscripts that have scientific merit but also methodological problems that are judged to be rectifiable by supplying more information about procedures, clarifying points, or conducting additional data analyses; (3) a combination of 1 and 2, which is probably the most common category.

One relevant question from an author's perspective is: What are the chances for my manuscript to be accepted? The answer to this question depends on the individual journal and its rate of submissions relative to available journal space. Readers might wish to consult the 1996 Summary Report of Journal Operations for detailed information concerning the rejection rates of a wide range of journals (American Psychological Association, 1997). Readers may be interested to know that from 1992 to 1997, JPP's rejection rate averaged 80% (LaGreca, 1998).

The Editor's Decision Letter

From an author's perspective, the most critical communication is the editor's decision letter, which reflects a synthesis of the reviewer's and editor's comments and contains a specific recommendation for what the author should do, for example, revise the manuscript or submit it to another journal. Authors also receive the narratives of individual reviewers' critiques but do not receive their specific recommendation for acceptance or rejection or comments to the editor.

Most of the time an editor's letter will provide information that clearly describes the nature of and reasons for his or her decision. Nevertheless, speaking from my own experience as a rejected author, it is not always easy to tell how or why an editorial decision was made based on the information that is given in the decision letter. The author's task in deciphering the rationale for a negative editorial decision is especially difficult when comments from one or more of the reviewers convey a high level of enthusiasm for the manuscript. Authors need to understand that it is the editor who makes the final judgment concerning the manuscript's acceptance. Not surprisingly, an editor may be persuaded by some reviewers' evaluations more than others may. Competent editors consult with and listen to their reviewers but make their own decisions concerning the disposition of a manuscript.

HOW DO REVIEWERS EVALUATE THE OVERALL QUALITY OF AN EMPIRICAL ARTICLE?

The decision concerning the acceptance and/or rejection of an empirical research report is a cumulative global judgment concerning the quality, merit, and scientific contribution that is made by the study. What do the reviewers and editors consider in making such judgments? Authors should recognize that individual reviewers may weigh various elements of a manuscript differently and may have different thresholds of concern about specific methodological issues (Fiske & Fogg, 1990; Peters & Ceci, 1985). Nevertheless, reviewers generally base their judgments on the following areas of the manuscript: (1) relevance of the study to the journal's mission; (2) clarity and significance of the scientific question; (3) new knowledge contributed; (4) how the scientific question was tested (study design and data analysis); (5) quality and strength of the findings and quality of their presentation; and (6) quality of the author's discussion of the study's implications and limitations. Each area is now described.

Is the Study Relevant to the Journal's Mission or Purpose?

Studies that are judged to be tangential to the central mission of the journal will either be rejected outright or rated lower than those that are judged to be more centrally relevant. Authors should recognize that even high-quality manuscripts will not be accepted if they are judged to be irrelevant to the journals' mission. In my experience, manuscripts that contain research topics that cut across different fields (e.g., child development and pediatric psychology) may make it difficult for reviewers to judge the relevance of a manuscript to a journal's mission. Conse-

quently, the “take-home” message here is that authors should make sure that their research fits the journal’s mission before they submit the manuscript for publication. Moreover, authors should take some care to ensure that relevance of their work to the journal’s mission is clearly stated rather than implicit, most especially if their research cuts across different fields. When authors are not clear whether or how well their research fits a journal’s mission, it is not only permissible but also desirable for them to contact the editor directly to clarify this issue.

Is the Scientific Question Clearly Stated?

Reviewers and editors are interested in determining the answer to one basic but surprisingly elusive question about each manuscript: What is the basic scientific question that is addressed? It is usually surprising to authors (I know it has been for me!) to learn that reviewers are sometimes confused about the essential question that they posed in their research. Authors should recognize that it is easy to be so immersed in their research, which is generally both familiar and significant to them, that they may fail to clearly articulate the scientific question to others who are not so familiar with the research topic. Clarification of the primary study question is a special challenge in complex research projects that address multiple, interrelated questions such as: What factors predict psychological outcomes including self-esteem and body image in multiple chronic illness groups, such as cystic fibrosis and diabetes?

What Is the Significance of the Study Question?

Editors are most interested in publishing studies that address scientifically significant questions that advance knowledge in the area of the research. Reviewers generally consider several questions in making a judgment about significance: Does the study address a question that is important or significant to the field? Does the manuscript address a novel question? Alternatively, if a study addresses a question or an issue that has been studied before, is the question answered in a more definitive or methodologically sound way than in previous research?

The difficult subjective judgment concerning whether a study addresses a significant question is one that rests on expert consensus. Consequently, it is not always easy for authors to ascertain what is a significant scientific question unless they have extensive research experience in and familiarity with the topic of interest. For this reason, consultation with colleagues and careful reviews of the state of the art of research in one’s field of interest can provide information to support the significance of a topic.

Research can have different types of significance. For example, an author may claim *theoretical* significance for a study that advances or tests a theory. Alternatively, an author may emphasize the *methodological* significance of research that addresses a question more carefully or with different methods than in previous research. Finally, an author may also claim *clinical* significance if their research contributes to the development of a new method of assessment, treatment, or contains new information concerning treatment efficacy.

Do the Data Contribute New Scientific Knowledge?

As is the case for questions of significance, the judgment about whether research findings make a novel contribution to scientific knowledge is a difficult one that depends on the level of development of science in a given area of research at a particular time. A novel, exciting research question and data set at one time point may emerge as less compelling as a field evolves. For example, in the early 1970s, very little was known about the psychological adjustment of children with chronic physical illness. At that time, descriptive, controlled research of most chronic illness groups provided new scientific information. Now, decades later, the wealth of available descriptive research concerning the psychological adjustment of children with different chronic illnesses (Thompson, & Gustafson, 1996) had placed a burden on researchers to advance scientific knowledge beyond simple description by conducting research that advances knowledge of processes that contribute to psychological adaptation (Drotar, 1997a) or evaluates the efficacy of interventions with this population (Drotar, 1997b).

The author's ability to convince reviewers that his or her research makes a new as well as significant contribution is a key factor in an article's eventual acceptance. In some cases, an author's findings may be both novel and significant, but the presentation of the work may not effectively establish these points in the minds of reviewers. Sternberg (1988) has observed that authors often make mistakes in "selling the wrong product," that is, failing to articulate the essential new contribution of their work or by not highlighting it effectively. Authors need to clearly state the essential new contribution of their research in the introduction to their work and underscore it again in their discussion. In doing so, authors must steer a difficult line between underselling the contribution of their research versus overstating it and claiming too much (Sternberg, 1988).

How Is the Study Question Answered?

Editors and reviewers generally distinguish between the significance of a study's scientific question versus how the question is addressed. Research can address a significant question and may have some interesting findings but not make a scientific contribution because of a faulty experimental design. For this reason, the methodological quality of a manuscript is a key factor in the eventual acceptance or rejection of an empirical journal article. Reviewers' and editors' judgments about the methodological quality of manuscripts are based on many factors including the clarity and quality in statement of hypotheses, study design, sampling and recruitment of subjects, analyses, measurement, data analyses, and presentation of findings. Issues related to basic research methods are beyond the scope of this chapter but can be found in basic texts (Breakwell, Hammond, & Fife-Shaw, 1995; Creswell, 1998; Kazdin, 1980, 1992; Rosenthal & Rosnow, 1991).

Are the Findings Strong and Clearly Stated?

There is no way around it: An author who is fortunate enough to garner powerful findings, especially those in accord with his or her hypotheses, clearly has an edge over his or her counterpart who has designed an elegant study that has

demonstrated no statistically significant effects. The burden of proof is so difficult for investigators who do not find significant differences because multiple factors can lead to an insensitive test of hypotheses and contribute to nonsignificant findings. Nevertheless, elegantly designed and well-executed tests of important clinical and scientific questions may sometimes be discriminated against for publication (Easterbrook, Berlin, Gopalan, & Matthews, 1991). The problem of publication bias has led to cogent arguments to consider effect size and confidence intervals rather than statistical significance in reporting and evaluating data (Rosnow & Rosenthal, 1989). Some editors will publish findings that are not statistically significant provided that authors can make a convincing case that their study was well executed and the findings contribute new or important knowledge.

In other instances, authors may address an important scientific question with a well-designed study and data analytic plan, but their findings may not be clearly presented. It is quite possible for authors to bury the importance of their findings in a disorganized, ambiguous presentation of their results and/or by presenting non-contributory post hoc analyses that only serve to distract the reader from the most important data that is presented.

Are the Implications of a Study Novel and Clear?

An author's ability to translate his or her study's findings into a clear, take-home message of the study is an important but elusive commodity (Sternberg, 1991). For maximum impact, the study's scientific contribution should be clearly stated at the outset and described in discussion together with the implications for future research (Sternberg & Gordeeva, 1996). Also, authors should carefully craft their discussion sections so that relevant issues that need to be considered in interpreting their findings, such as methodological limitations, are stated but not overdrawn.

PRESENTING THE DETAILS OF A RESEARCH REPORT: WHAT DO REVIEWERS LOOK FOR IN THE SECTIONS OF A MANUSCRIPT?

The following section summarizes the specific items that reviewers look for in evaluating the various sections of a research report and implications for authors. To guide their reading of this section, readers should consult Table 1, which summarizes feedback to authors that was given consistently by reviewers for JPP (LaGreca, 1994a), and Table 2, which is an authors' checklist that contains useful reminders (LaGreca 1994b).

Introduction

Reviewers look for a clear statement of the purpose and goals of research. Consequently, the significance of the study and its relationship to previous research should be carefully described (see previous section). Moreover, the author should provide a clear rationale for the primary variables that are studied and their relationship to the primary research question and study goals.

Given the relevance of hypotheses for guiding data analyses and in evaluating

Table 1. Feedback Checklist for Submitted Manuscripts^a

INTRODUCTION
—Greater clarity is needed with respect to the conceptual or theoretical rationale for the study and the study goals/hypotheses. (It should be clear from this section what the major contributions of the study are and how it extends previous research in this area.)
—The linkages between the study goals and variables assessed are not clear. The introduction should set up a clear rationale for examining the variables that are measured and analyzed.
—Other relevant literature should be cited if the work clearly relates your work to prior research.
—The study's main contributions are not clear (e.g., what new scientific information is provided beyond existing work?).
—The study lacks a consistent focus. Structure the paper clearly around a primary goal/theme and/or central question.
—Hypotheses should be stated clearly, unless the study is explicitly exploratory in nature.
METHOD
Subjects
—Provide information concerning the initial pool of eligible subjects (e.g., all patients with asthma followed at a pediatric center).
—Include details regarding the subject sample (e.g., age, gender, race, and ethnicity). Other demographic variables of interest (family composition, socioeconomic status, parental education, etc.) may also be needed.
—Provide details on how subjects were selected/recruited. Explain and provide rationale for any exclusionary criteria.
—Report subject participation rates, compare subjects who refused versus agreed to participate.
—When multiple groups are used (e.g., intervention and control groups; different disease and comparison groups), report demographic data separately for each group, preferably in a table. Report results of <i>t</i> -tests, χ^2 analyses, etc., comparing the groups in a table or the text.
—When matching subjects for age, or other variables, the decision rule should be explicit (e.g., if subjects are matched on age, are they within 6 months, plus or minus, of the target subject?)
Procedure
—This section should be separate from Subjects or Measures sections. Explain how subjects were recruited, how the measures were administered and to whom (mother, primary caretaker, both parents, child, etc.), who the experimenters were, and where the procedures were conducted (e.g., home, office, laboratory).
—Provide details on how informed consent from parents was obtained, as well as child assent.
—Explain treatment procedures in detail. How was the intervention conducted and by whom? How often was it administered and how long were the sessions? The rationale for the various treatment elements should be explicit. Others should be able to clearly follow if not replicate your procedures from reading this section.
—Information on the availability of treatment manuals should be indicated in a footnote.
Measures
—Describe each measure. Measures should be clearly related to the study hypotheses/aims.
—Details on scoring, coding, reliability, and validity should be included. Indicate which scores are used in the subsequent analyses. Unless the instrument is very widely known (e.g., Child Behavior Check List), details should be provided on the reliability and validity of the measure.
—Avoid the use of single item measures. If this is not possible, use caution throughout the text and interpret the results cautiously.
—Make sure that all measures included clearly fit with the described study goals and hypotheses. If not, this may detract from the overall presentation of the study and contribute to lack of clarity in the paper's focus.
—If a revision of an established measure is used, provide data to support the revision. For example, the internal consistency of a revised measure cannot be assumed to be the same as the original.
—For newly developed measures, provide details on how the items were developed, how many items there were, how the final pool of items was derived, how the items and scales were scored, what the possible range of scores was, and which scores (mean? total?) were used in data analyses.

(continued)

Table 1. (Continued)

-
- When unconventional or preliminary measures are used, greater caution and more caveats need to be integrated throughout the text and discussion.
 - When measures are administered to more than one individual (e.g., both parents), explain how the multiple measures are handled in the analyses.
 - For intervention studies, measures of treatment integrity should be included, to indicate that the intervention was administered as planned. Otherwise, this must be discussed as a study limitation.

RESULTS

Reporting Mean, Standard Deviations, Significance Tests

- Include a listing of means and standard deviations (with accompanying sample sizes) for all measures (even when the analyses are not significant) preferably in table form. If multiple groups are used (e.g., children with diabetes, cystic fibrosis, etc.), these data should be provided with each group.
- Results should be organized around the issues/questions posed in the Introduction.
- Report the results of statistical tests (F value, degrees of freedom, etc.). Check the APA manual and/or recent issues of the journal.
- Results should not be reported/discussed as significant if alpha levels exceed .05.
- In tables and in the text, use “conventional” alpha levels (e.g., $P < .05, .01$) rather than “exact” alpha levels (.018).
- Statistical power for analyses should be described and justified.
- The one exception to the above is when adjustments are made to the alpha level to minimize experiment-wise error (e.g., Bonferroni corrections). In such cases, indicate what the revised alpha is, and use this in the description of significant results.

Multivariate/Regression Analyses

- The use of multiple dependent measures requires multivariate analyses (e.g., MANOVAs) or at least a correction of alpha levels for type I error (e.g., Bonferroni correction).
- The basis for the hierarchical regression model tested should be explicitly foreshadowed in the Introduction, ideally with specific hypotheses, and explained in the Results section.
- The rationale for the regression analyses, the variables included in the analyses, and the presentation of these analyses should be clearly communicated.
- When presenting regression analyses, show each step in the model, listing variables entered on each step, and the associated R^2 , F tests (for each step as well as overall F), and significance levels. Standardized beta coefficients should be included, along with their significance levels.
- Demographic variables (e.g., gender, ethnicity) need to be considered in the analyses as covariates.

Data Reduction

- Are all variables equally important to include? If so, a clear rationale should be provided for these variables.
- Are the dependent variables strongly interrelated? If so, multivariate analyses or other efforts to limit type I error would be desirable.
- Are the independent variables in a regression analysis strongly interrelated? If so, consider data reduction or choice of a particular “representative” variable.
- When using multiple measures of a construct, efforts should be made to reduce or consolidate the data. If measures are highly interrelated, can they be combined in some manner to obtain an overall index? The use of Structural Equation Modeling may be helpful in some instances.

Use of Current Test Norms

- The analyses that pertain to the Child Behavior Check List (CBCL) need to be redone using the 1991 norms (Achenbach, 1991).
- The analyses of _____ need to be redone using current norms.

Clinical Significance and “Caseness”

- Indicate whether and how statistical differences were also clinically significant.
- Present data on subjects who represent clinical cases (e.g., scores about the 98th percentile on CBCL). Justify the procedure for “caseness” and compare with others in the literature.

Age/Developmental Issues

- Subjects’ age may affect the interpretation of your results. Analyze data with greater sensitivity to developmental issues.
-

Table 1. (Continued)**DISCUSSION**

- Do not restate the results unnecessarily. At/near the beginning of the Discussion, highlight the main points of the paper, indicating how the results support your position.
- Do not introduce new results in this section.
- Discussion should be clearly related to the issues discussed and the questions posed in the Introduction.
- Discussion should be succinct and focused on the positive and negative findings, relating them to previous research. Address whether the main study questions were supported or not; do not feel obligated to discuss each and every finding.
- Note competing explanations for findings and discrepancies with previous research.
- Discuss results that did not turn out as expected.
- Address the specific implications of your findings for pediatric psychologists.
- Include a thoughtful, detailed discussion of the study's limitations, especially factors that might limit the inferences that can be drawn.
- Greater caution is needed in interpreting the study findings; make sure that the conclusions do not go beyond the results.
- Discussion suggestions for further research, especially the "next steps" that are suggested by your work.

RELEVANCE TO PEDIATRIC PSYCHOLOGY

- The study's relevance to pediatric psychology should be explicit, especially in the Introduction and Discussion. For example, discuss the clinical implications of the findings for pediatric psychologists; include references of a pediatric nature (e.g., from *JPP* and other pediatric sources), etc.

^aFrom LaGreca (1994a).

findings, reviewers also look for clearly stated, specific hypotheses in the introductions of research reports. Specific hypotheses also serve as important guides to help reviewers to understand how the author's primary research question was operationalized and tested. Reviewers appreciate manuscripts in which they can clearly follow the rationale for hypotheses that are based on a clear conceptual framework, theory, and/or empirical research.

Describing Methods

In order to judge whether authors conducted an effective test of their study questions, reviewers need to understand the basic research design that was used. Consequently, they want crystal-clear answers to the following questions. How many groups were in the study? Is this a correlational or experimental design? What is the rationale for the design that was used? Is the design clearly related to the research question? To ensure that reviewers can find the answers to these basic questions, it is useful for authors to include a brief statement of their design in the method, for example, "this was a randomized, controlled study comparing the outcomes of two groups of children with anxiety disorders, one that received cognitive-behavioral treatment and a wait-list control that received no treatment."

To make sure that the rationale for their study design is clear to reviewers, authors also may want to indicate this in a statement such as the following: "The comparison groups for a study of psychological adaptation of children with cystic fibrosis included a group of physically healthy children of comparable age and

Table 2. Formatting Checklist for Submitted Manuscripts^a

TITLE PAGE (Page 1)
—Include: Title, Authors, Affiliations, and Address of Corresponding Author.
—Type a running head (centered) at the bottom of the page and use it throughout the paper, in the upper righthand corner, along with page numbers.
—Acknowledgments should be typed on a separate sheet, and placed <i>before</i> the title page, to facilitate removal for blind review. Author notes and names should <i>not</i> appear anywhere in the text.
ABSTRACT (Page 2)
—Begin the Abstract with a verb (current APA style).
—Do not exceed 1000 characters (including spaces) or approximately 100–120 words.
—Include Key Words (or phrases) near the bottom of the Abstract page.
INTRODUCTION
—Begin the introduction page with the title of the paper.
METHOD
General Issues
—Include separate sections for Subjects, Procedures, and Measures (see APA style manual).
RESULTS
Reporting Means, Standard Deviations, Significance Tests
—Include a listing of means and standard deviations (with accompanying sample sizes) for all measures (even when the analyses are not significant) preferably in table form. If multiple groups are used (e.g., children with diabetes, cystic fibrosis, etc.), these data should be provided for each group.
—Report the results of statistical tests (F value, degrees of freedom, etc.). Check the APA manual and/or recent issues of the journal.
—Results should not be reported/discussed as significant if alpha levels exceed .05.
REFERENCES
—References should match in name/date the citations in the text. Check correspondence between text and reference list.
—Use APA style in formatting references and in their citation in the text.
—Reference list is longer than needed.
TABLES
—Consolidate into one table data presented in Tables ____.
—Include the standard deviations, as well as the means, in Table ____.
—Include significance tests (t, F, etc.) and P-levels (.05, .01, .001) in Table ____.
—Delete from Results data that are presented in Table ____.
—Check APA style manual for format of tables.
WRITING STYLE AND PRESENTATION
—The manuscript is too long. Do not exceed 25 pages (14 pages for a <i>Brief Report</i>). This count includes the title page, references, tables, etc.), with one-inch or greater margins, and no more than 26 lines per page. Do not use small print. Begin numbering with the title page.
—Double-space throughout (e.g., Abstract, References, Tables).
—The tone of the terminology is more negative or pathological than is necessary.
—Use the convention of “person first” rather than “disease first” (e.g., children with diabetes, rather than diabetic children).
—Wording, spelling, or grammatical errors need correction.
—Check the wording carefully for redundancies.
—Too much psychological jargon is used; describe terms more simply.

^aFrom LaGreca (1994b).

socioeconomic status as well as a group of children with diabetes, which was recruited to control for the nonspecific effects of having a chronic physical illness.”

Clarifying Procedures for Sampling and Recruitment of Subjects

Reviewers want to see a description of the sample selection and recruitment procedures that are well-rationalized and sufficiently clear and detailed to allow replication by others (see Chapter 4, this volume). Consequently, investigators should describe their methods of sampling in some depth, including specifying the initial pool of subjects [e.g., participants were selected among “children with attention deficit disorder with hyperactivity (ages 6–12) who were referred to pediatricians in a particular clinic setting”]. It is also useful to describe the rationale for selection of the groups, unless this is clearly stated in the introduction. Investigators also should describe how their subjects were selected and recruited for the study, for example, “Patients between the ages of 11 and 17 with diabetes of 1 year’s duration who were receiving medical care at a clinic center were identified through review of the patient roster. The families of eligible patients were then informed of the study by their physician and contacted by the investigator over the phone to determine their interest in participation.” The rationale for the criteria that were used to select subjects should be specified, because these are often difficult for reviewers to ascertain. For example, in the study described above: “children ages 11 to 17 were chosen because 1 year’s duration was considered necessary to define a condition that was chronic.”

In presenting information about the sample, investigators should describe the relevant demographic characteristics (e.g., age, gender, ethnicity, and socioeconomic status) in sufficient detail so that reviewers have a clear sense of who participated in the study. Most reviewers want to see a description of relevant demographic characteristics, such as family socioeconomic status, age, and/or gender, of the sample who did not participate, reasons for their nonparticipation (refusal, could not be contacted, etc.), and comparison of participants and nonparticipants on relevant demographic characteristics (see Chapter 4, this volume). If matching procedures are used, they should be clearly specified so that reviewers can follow the authors’ decisions concerning selection and implementation of the matching procedures.

Procedures

In writing the procedures section, investigators should put themselves in the position of someone who is beginning a research study and address the question: What would I need to know if I were to replicate this project? Reviewers want to see procedural information that is sufficiently detailed to allow them to follow what was done in the study (e.g., how procedures were conducted), how measures were administered, where (e.g., office setting, laboratory, home), and by whom (e.g., research assistants with specialized training) (see Table 1).

In presenting information about their procedures, most authors struggle with the difficult trade-off between the need to convey a sufficient level of information that is necessary to facilitate editorial evaluation and replication by others versus the need for a succinct presentation. Because they can always cut extraneous

information when they respond to reviews, authors may want to err on the side of completeness in describing their methods in their initial submission (within limits, of course).

Describing Procedures for Informed Consent

One important feature of procedures, which is overlooked in a surprising number of manuscripts, is how informed consent was obtained from parents and, if applicable, assent from children and adolescents (Range & Cotton, 1995). Information concerning the special considerations that were made to attend to ethical issues in specific populations that were studied, for example, referrals for intervention for individuals who were found to be depressed, also should be described (see Chapter 14, this volume).

Measures

Because reviewers often have many questions about measures (see Table 1), these should be clearly and carefully described along with details on scoring, coding, reliability, and validity. Ideally, information concerning reliability and validity should be presented for the sample that is studied rather than for the standardization sample. Investigators who revise established measures should provide data concerning their reliability and validity. If such data are unavailable, investigators should articulate the relevant qualifications in study conclusions. For newly developed measures, authors should describe how the items were developed, how the final pool of items was derived, and reliability and validity.

One common problem identified by reviewers is an author's failure to clarify the rationale for including a specific measure. Readers also should be apprised of how measures relate to the study hypotheses, as in the following example: "The Children's Depression Inventory was employed to test the hypothesis that children with newly diagnosed cancer would be more depressed than a comparison group of children whose cancer is in remission." It also can be helpful to clarify the choice of a specific measure, as in the following: "The Harter Self-Perception Scale (Harter, 1985) was used to provide a multidimensional assessment of self-esteem that limits the impact of social desirability on children's responses." Finally, when using more than one measure of the same construct, authors should give a rationale for this decision rather than assume that it is evident to the reviewers, for example, "in order to provide information concerning level of anxiety from the perspectives of different informants, two measures of anxiety were utilized: self-report and parent report."

Presenting Results

In my experience, reviewers frequently identify the results section of a manuscript as problematic or confusing (see Tables 1, 2; Squires, 1990). Authors should appreciate that they need to lead their readers through their results step by step, rather than take them on a long, winding journey without a compass. Such guidance can be provided in several ways: For example, it is generally helpful to begin with a brief sentence that informs readers about what will be presented in the

results: for example, general descriptive findings, tests of hypotheses, secondary findings. Readers also should be given a rationale for statistical tests that were conducted, for example, “to test the hypothesized differences between group X and group Y, an analysis of variance was performed,” or “Pearson product-moment correlations were performed to assess the relationship between variable *x* and variable *y*.” Finally, relevant findings should be highlighted and primary, hypothesis-driven results should be clearly distinguished from secondary analyses.

Reviewers like to see continuity between the introduction of a study and results. Consequently, it is helpful for authors to structure their findings around the major hypotheses and the questions that they raised in their introduction. Sub-headings also provide a useful way to highlight the major sections of the analyses and to distinguish between primary and secondary data analyses. Summary statements also can be helpful in clarifying the description of various findings. For example, statements such as, “children who reported higher self-esteem were rated as better adjusted by their parents,” can be used to describe a significant correlation between children’s reports of self-esteem and parents’ reports of their adjustment.

Tables and figures provide efficient, reader-friendly ways to summarize key findings (Wallgren, Wallgren, Persson, Jorner, & Haaland, 1996), but they should be used sparingly to present information that cannot easily be summarized in another way. Primary candidates for presentation in tables include the following: (1) descriptive characteristics of the sample; (2) means and standard deviations for major variables; and (3) summaries of hypothesis-driven analyses, for example, correlation matrices, analyses of variance, and regression analyses. While tables provide effective ways of presenting research findings, they do not substitute for describing the results in the text and guiding readers through the findings.

Information concerning statistical power is often neglected in presenting research findings but is critical to reviewers’ evaluation of whether the study hypotheses received an adequate test (Cohen, 1992). An increasing number of journals now require that effect sizes be reported to enable readers to clarify information concerning power, and authors should report this information (Gardner & Altman, 1986). Given the fact that increasingly specialized methods of data analysis are used in research in pediatric and clinical child psychology, authors should be alert to the special problems that arise in presenting information from various statistical analyses. See Table 1 for specific examples.

Discussing the Scientific Contribution of One’s Research

In many ways, the discussion section of a research-based manuscript is the most difficult to write. Common problems in preparing the discussion include restating the results and discussing each and every finding, which makes for a dull, unwieldy presentation (Bem, 1987). In contrast, readers appreciate a dynamic summary of the essential contributions of the present study, along with how it fits with or is discrepant from previous research findings, theoretical models, and so forth. To create a more interesting discussion, each finding should not be reconciled with previous research findings. Moreover, the discussion should be coherent with the manuscript’s introduction so that readers can clearly follow the confirmation or disconfirmation of salient hypotheses (Bem, 1987).

Readers of discussion sections also appreciate authors' attention to the following questions: What next steps in research are suggested by the findings? What critical questions need to be addressed to advance the field? In writing the implications, it is important for authors to be as creative yet as specific as is possible.

Other important payoffs for readers include discussion of clinically relevant implications of the findings and proposed directions for future research. In particular, clinical or policy implications have been neglected in articles in pediatric and clinical child psychology but are often extremely helpful, as well as interesting (Roberts et al., 1992). On the other hand, to maintain credibility with reviewers, authors need to be very careful in stretching clinical implications (or other conclusions) far beyond their data.

No piece of scientific research is without flaws. Consequently, it is important for authors to describe most significant limitations of their study in order to help readers understand them and judge their potential influence. Authors need to demonstrate that they recognize the major limitations of their research rather than leave it to their readers to identify them. In presenting the limitations of their work, authors also should consider their specific impact, for example, how did the problem affect internal validity or external validity (generalizability) of findings (Campbell & Stanley, 1963)? Finally, it may be useful for authors to inform readers about limitations of their studies that can or should be addressed in future research. In some instances, addressing methodological limitations may launch exciting new research directions.

GENERAL GUIDELINES IN WRITING RESEARCH ARTICLES

Enhancing Clarity Of Presentation

Because reviewers are not necessarily familiar with the major content areas or issues presented in manuscripts, clarity of writing is a key concern for authors. One way to enhance the clarity of one's manuscript is to develop an outline in which the essential points in a manuscript are listed and then to go back over the manuscript (and/or have a colleague do this) to make sure that all the points or themes that are listed are clearly and completely articulated.

Integrating Components of the Manuscript

To be most effective, a manuscript should form a coherent, integrated product. The various components of the manuscript should not only be consistent with one another but where possible refer to and connect with one another. For example, the theoretical or research issues raised in the introduction can be elaborated in the discussion in light of the findings. Hypotheses that are presented in the introduction should guide the data analysis and presentation of the results.

Proofing the Manuscript

Reviewers will notice a manuscript that is riddled with errors and inconsistencies, but such negative attention is not what authors desire. Consequently,

authors should take primary responsibility for proofing their manuscript to catch typographical errors and obvious problems in sentence structure and to review references for completeness. Because it is difficult to catch all of one's own writing mistakes, especially when one has worked continually on a manuscript, it is useful to have colleagues or trusted assistants carefully review the manuscript for errors.

REVISING SCIENTIFIC MANUSCRIPTS

Successful authors learn to accept that it is commonplace for manuscripts to go through one or more revisions before being eventually accepted. In fact, given the high rejection rates of many journals, authors should regard a reasonably encouraging revise-and-resubmit editorial verdict as a triumph. Nevertheless, revise-and-resubmit decisions and rejections are difficult to accept from an emotional standpoint. Speaking from experience, no one likes to see all the problems of one's scientific writing and research critiqued in black and white, especially by three to five experienced scientists, some of whom may be even be friends and/or colleagues. Because novice authors have little experience in receiving detailed critiques, they may be particularly troubled by negative reviews (Roberts et al., 1982).

In this regard, I remember how I felt when I received the reviews of my first senior-authored manuscript submission, which was based on my master's thesis. Being new to the world of scientific publishing, I did not anticipate problems in the review. After all, my research had already been through the careful scrutiny of my committee. Consequently, I was shocked when my work received much less than the positive reception I had in retrospect foolishly expected and was rejected. It took about 6 months for me to work with the manuscript once again in order to determine what I should do next (I resubmitted elsewhere and the paper was eventually accepted by another journal). Twenty-five years later, I still do not look forward to reviews but have grown accustomed to the experience. When faced with a revise-and-resubmit or reject editorial verdict, I now try to follow Roberts et al. (1982) useful axiom: "read, react, revise, and resubmit" (p. 38).

Deciding Whether and Where to Resubmit a Manuscript

The first issue for authors to decide is whether they want to take on the work of revising their manuscript (Nagata & Trierweiler, 1996). In my experience, authors generally weigh several cost-benefit considerations in making the decision to revise and resubmit their manuscript to the original journal versus sending the manuscript to a different one. These include: (1) their interpretation of the editor's interest and appraisal of their manuscript; (2) the importance they attach to publishing in the initial journal versus another outlet; (3) their appraisal of the amount of work that is required in the revision; (4) their judgment about whether it is possible to make the suggested changes; (5) how important it is to them to have the article published; (6) competing demands that would interfere with a timely resubmission; and (7) feasibility of responding to the critique effectively. For example, when reviewers ask for new data collection and changes in the design, it may not be feasible to respond to their critique. Even if revisions are feasible to make, authors may decide that the work that would be required by the revision is simply not worth

it to them, especially in light of the fact that the eventual editorial outcome is not guaranteed.

Strategies to Enhance Acceptance of a Resubmission

Assuming that one has made a decision to resubmit the article, several strategies can be used to enhance the likelihood of eventual acceptance (Nagata & Trierweiler, 1996). An author's first task is to read and reread the critiques to make sure they are understood. In some cases, reviewer's comments may be sufficiently vague or open-ended that it is not clear what is needed in a revision. In view of the importance of responding to each reviewer's comment in detail, authors need to rely on their coauthors, colleagues, and if need be a discussion with the editor to make sure they understand what is meant by the various points in the critique (see Chapter 19, this volume).

In other situations, the reviewers' critiques may be reasonably clear, but the best strategy to manage them is not at all obvious. Such confusion is understandable because there are alternative approaches to manage problems that reviewers raise, for example, approaches to data analysis. For this reason, discussion with colleagues can be extremely helpful in deciding how best to manage reviewers' suggestions.

Before beginning a revision, I have found it helpful to put the reviews aside briefly (but not 6 months, as in my initial experience) to let my initial reactions wear off and to obtain some perspective. However, as in solving other problems, avoidance is not the best long-term strategy in managing a revision. I also have found it helpful to begin the process of revision by composing a draft letter of response to the editor, because this helps me to organize my thoughts and begin outlining the work that needs to be done in the revision (Nagata & Trierweiler, 1996). Most journals require authors to develop a detailed, point-by-point response to the critique to accompany the resubmitted manuscript that clarifies and highlights revisions and/or provides a vehicle for rebuttal, in the case of disagreement about methods. I also have found it helpful to begin by first making the easiest changes, such as straightforward editorial suggestions, while I am pondering my options to manage the more difficult problems.

Responding to Reviewers' Concerns

In responding to reviews and composing the letter to accompany the revised manuscript, it is important to attend carefully to each and every point that was noted by the reviewers and/or managing editor, paying particular attention to the editor's critique but neglecting no one's points. Irrespective of whether one agrees with specific points of the critique, a thorough response to the editor's letter and reviewers' comments in the cover letter can make a difference in the eventual editorial verdict. This is because an author's responsiveness to critique is carefully considered by reviewers.

On the other hand, while a respectful, point-by-point response to a critique is always indicated, authors need not slavishly capitulate to every editorial concern. In some instances, authors may have valid disagreements with reviewers about the importance of specific methodological issues and may want to make a cogent

argument for their point of view by preparing a detailed rebuttal. The editor will then decide whether the author or reviewer's point is most valid. Such rebuttals are best done in the context of the authors' detailed cover letter, and for maximum efficacy should have a factual but not condescending or argumentative tone. The authors need to acknowledge that they heard the point of criticism, have thought carefully about it, but have a different point of view that is informed by scientific evidence and/or experience. The tone of the letter of rebuttal should be one of: "I respectfully disagree," rather than "the ignorant reviewer was dead wrong" (even though this sometimes may be the case!).

In their letter to the editor that describes their revision, authors should make it crystal clear what recommended changes they made, how they were made, and where in the manuscript they were made. If recommended changes were not made, authors should indicate why they did not make them. In many cases, the editor will send a manuscript back to the reviewers who will undoubtedly want to see how their points were heard and managed. For this reason, it is best for authors to guide their audience through this process by repeating or paraphrasing each specific point that was made by the reviewers and editor, to describe their response to the point, and highlight their responses.

Submitting a Revised Manuscript to a Different Journal

Authors who receive a revise-and-resubmit or rejection editorial verdict from one journal are free to submit their manuscripts to another one. When the reviewers' points cannot be addressed and/or when the time and energy involved in responding to them is not viewed as warranted, submission to another journal may be the best strategy. One question that arises in such resubmissions is how much of the revisions that are requested by one set of reviewers should be included in a submission to another journal? In my experience, it is wise practice to regard the review as a consultation and to do whatever can be done to improve the quality of the manuscript before sending it out again. Nevertheless, when balancing the manuscript's chances of acceptance against the time, effort, and feasibility of making the revisions, some authors will choose to make some but not all of the suggested revisions. Others may select a minimalist strategy by sending in essentially the same manuscript for a second opinion. One problem with this latter strategy is that there is a chance that the same reviewers as in the initial submission will review the manuscript. Members of editorial boards in child clinical and pediatric psychology journals overlap and there is a small pool of experts for some research topics.

MANAGING DIFFICULT ISSUES IN PREPARING ARTICLES FOR PUBLICATION

Managing the Process of Writing with Coauthors

One of the more challenging and interesting issues in writing empirical articles is working with coauthors. On the one hand, coauthors can be an invigorating and important source of intellectual energy. On the other, collaborative writing, espe-

cially with those who have a very different style from one's own, can be a source of conflict and tension. Given the need for organization and leadership to bring a manuscript along to completion, my experience suggests that it is most effective to decide on a lead author and the order of authorship credit very early in the process. Lead authors should decide how best to involve their coauthors. For example, in some situations, it may be most effective to divide the responsibility for different sections. If this strategy is chosen, I would suggest that authors elect one person to assemble the components of the manuscript, lest the manuscript resemble a less than artistic patchwork quilt style. Alternatively, the lead author may decide to put together a draft of the entire manuscript and have coauthors critique the work and add specific sections.

It is also important for a lead author to understand the work styles of coauthors and communicate with them about their specific roles and responsibilities in the writing. For this reason it may be useful to develop a shared outline of the article and agree on a time line. I have found that the most difficult problems are coauthors who do not complete their responsibilities in a timely manner. It takes diplomacy and sometimes firm limit setting to manage writing with coauthors.

Potential Conflicts with Coauthors

One of the most challenging problems in preparing publications that are shared among different collaborators is raised by decisions concerning authorship credit (see Chapter 13, this volume). Conflicts around decisions concerning authorship credit include: (1) whether research warrants credit for authorship or acknowledgment, or (2) the order of the authors. According to the APA's (1994) guidelines, authorship is warranted for formulation of a research problem, hypotheses, or experimental design for organizing and conducting statistical analyses or for writing a major portion of the paper. Similarly, the International Committee of Medical Editors' guidelines for authorship credit include participation in the research design, writing, or revising of publications (Huth, 1986).

Irrespective of professional guidelines, prospective coauthors may have their own strongly held personal beliefs concerning authorship credit, which can raise ethical problems (Keith-Spiegel & Koocher, 1985). Collaborative, interdisciplinary writing can also generate conflicts about the dissemination of findings, especially the choice of publication outlet (Drotar, 1989; Taylor, 1987). Researchers who prefer to publish their findings in a journal that brings recognition and credit from their professional peers may regard publication in another profession's journal as less than optimal. Although there is no ready formula for making such difficult decisions, some collaborators may choose to compromise by selecting a professionally neutral interdisciplinary publication, alternating publications in different professional journals, or deciding on the best fit for publication of their paper based on the research content and/or journal circulation.

PUBLISHING MULTIPLE JOURNAL ARTICLES FROM A SINGLE DATA SET

Another difficult set of issues is raised by the decision to publish multiple papers from a single data set. Given the complexity of research projects in pediatric

and clinical child psychology and the availability of secondary data sets, authors' decisions concerning how many articles to publish from a single data set have become increasingly complex. Authors who work with complex data sets face a difficult task in balancing competing pressures to include all their data into one manuscript versus the need to restrict the scope of the information that is presented in any one article so that it can be managed by readers. Whenever it is feasible to do so, it is clearly most desirable to publish one comprehensive article from a large data set that is presented in an integrated form (Fine & Kurdek, 1994). Nevertheless, because the scope of research findings cannot always be integrated meaningfully in a single article, authors need guidelines concerning when it is appropriate to publish multiple journal articles from a single data set.

Fine and Kurdek (1994) have suggested that a reason for publishing multiple articles from a single data set is because an article that integrates the data is not comprehensible or possible. In order to determine whether these criteria are met, an author may attempt to write a single integrated article and have colleagues judge whether the article makes sense. Authors also may obtain guidance from editors concerning this issue by writing a single comprehensive article and then obtaining feedback.

Another criterion for publishing multiple articles from a single data set is when the secondary article(s) has a purpose that is clearly different from the primary article. This condition can be met if an article: (1) addresses different research questions; (2) uses different relevant literatures; and/or (3) includes different data pertaining to the different scientific questions that are addressed (Fine & Kurdek, 1994).

Longitudinal studies present unique problems for decisions about multiple versus single publications. Multiple publications of longitudinal data may be useful and necessary when: (1) longitudinal studies are designed to address several distinct purposes; (2) early results of longitudinal studies contribute new scientific knowledge; and (3) publications from different phases of the study are based on different subsets of the sample (Fine & Kurdek, 1994).

In the event that they choose to submit multiple articles from a single data set, authors should cross-reference the multiple articles based on their data, inform the editor about their plan to submit several papers, and submit previous papers with the current manuscript under review. In the case of multiple publications, detailed descriptions of the methods should be presented in the first article to which readers can be referred.

TRAINING IN SCIENTIFIC WRITING: ROLE OF A WRITERS' WORKSHOP SEMINAR

One of the most central yet difficult tasks for young researchers is to learn how to prepare their research for publication. It takes a great deal of supervised practice to communicate effectively in writing as well as feedback to recognize and address one's deficiencies. Unfortunately, research mentors who work with graduate students may not be able to devote sufficient time to facilitate the multifaceted skills that are necessary to write effectively in a wide range of formats (e.g., research reviews, grant writing, and research-based articles). In order to become proficient

in these different types of writing, students (and faculty) benefit from didactic training and practice in each of them. Moreover, the complex task requirements involved in scientific writing such as attention to deadlines, managing evaluation and constructive criticism, and dealing with multiple deadlines also require more time and attention in both didactic courses and mentorship.

Finally, traditional mentorship may not provide the most optimal context in which to teach students about the process of scholarly writing for several reasons: Some students are intimidated by their mentors who are much more experienced and knowledgeable about writing. Consequently, they may exaggerate the differences between their skills and that of their mentors. Students who personalize criticisms of their work may feel (at times erroneously) that they have special problems in producing writing and/or making their points clearly. In my experience, it is difficult for writers, including students, to appreciate that many of their personal reactions to writing and technical difficulties in writing are experienced by every author. While empathic mentors will communicate these issues to students, students benefit from opportunities to share some of the difficulties in writing and receive support for writing from their peers. For these reasons, I have found group-oriented methods of training in writing to be useful to help forge students' identities as writers and to give them practice in writing to facilitate their skills, as well as support to manage the inevitable difficulties associated with writing. One such method is a writers' workshop seminar which is described below.

Structure of the Writers' Workshop Seminar

I developed the writer's workshop seminar to enhance the writing productivity of a group of graduate students who were enrolled in a specialized training program that has been funded for the past 14 years by the National Institute of Mental Health to enhance the career development of researchers in the field of pediatric psychology (Drotar, 1998). Between four and six students are involved in this seminar which includes two basic elements: (1) a didactic component that includes lectures, review papers, scientific research articles, case reports, as well as an overview of the process of review and lectures on components of the manuscript, and (2) an experiential component that includes in-class review of students' writing projects.

The experiential component of the seminar, which is the most critical, is structured as follows: First, students select one or more writing project that they will work on for the class during the semester. A diverse set of writing projects have been reviewed in the seminar including manuscripts being prepared or revised for publication, masters and dissertation research proposals, research review papers, or grant proposals. Students are given the option to choose writing projects that are most relevant to their interests, professional development, and salient deadlines within their training program.

To set an agenda for the next class, several students agree to make "some progress" in their writing projects in time to distribute them to their instructor and classmates 1 week prior to the class to allow sufficient time for review. Depending on the length of the written material, two to three writing projects can be discussed in a 90-minute class. The definition of "progress in writing" is a deliberately

nonstringent one for several reasons: The seminar is designed to promote the notion that skill in writing is best achieved through a continuous, step-by-step process rather than by a “get it done in one sitting/back against the wall” strategy. For this reason, students are encouraged to submit first or second drafts of a project and/or “beginning stabs,” for example, a several-page draft of an introduction or an outline for a manuscript, and not be overly concerned about submitting lengthy or highly polished writing. In this way, students obtain firsthand experience with the principle that all writing is imperfect and can be improved with multiple drafts (Becker, 1986; Peterson, 1996). Students also are given choices concerning the timing and frequency of their presentations. Those students who have imminent deadlines and/or special zeal to finish a project may choose to present ahead of other students. However, all students are required to present their work throughout the semester.

Facilitating Learning about Writing through Group Critique

The core of the writer’s workshop seminar is the critique of students’ writing projects that is conducted by the instructor and students’ peers. The instructor moderates and leads the discussion of students’ writing projects. Students are encouraged to be honest in their critiques and also to contribute constructive comments that will be helpful to their colleagues in subsequent revisions. Suggestions to participating writers are given in the spirit of providing “consultations” rather than criticisms. Consultation is encouraged by informing students to guide their colleagues’ critiques by describing some of the difficulties they have encountered in writing and to identify the specific areas of the manuscript in which they feel they need the most help.

The group’s critiques tend to be free-ranging, including everything from suggestions for rephrasing text to more substantive reworkings of the manuscript. Whenever possible, the group critiques are used to illustrate common writing dilemmas (e.g., the difficulty of deciding how much detail to include in descriptions of previous research, assumptions that one’s writing is clear to others, etc.). A repeated message to students is that less than perfect is part of the territory for writers irrespective of their experience. This message can be communicated more emphatically if students also have a chance to critique their mentor’s writing.

Following the group critique, students are expected to make changes in their work within a specified period of time (usually 1–2 months or sooner if there is a specific deadline) and bring their work back to the group for review. Students who participate in the seminar are not required to follow each and every suggestion that is made by their colleagues. Rather, an emphasis is placed on a problem-solving approach in which students are taught to weigh the costs and benefits of incorporating various suggestions into their revisions. Moreover, students are encouraged to persist in their writing projects even if they are frustrated with or somewhat stalled in their writing and can manage only a few pages.

Finally, to enhance the efficiency in their writing, students are encouraged to work on multiple writing projects at once rather than to finish them, one by one, in a lockstep fashion. The experience of managing multiple writing projects gives students a more realistic taste of professional writing demands and the realities of subsequent professional careers.

Advantages of the Writers' Workshop Seminar Format

Based on discussions with students, perhaps the most important advantage of the writer's workshop seminar format is the repeated opportunities it affords students to present their written work and receive critiques from their peers and instructor in a relatively nonthreatening format. In this way, students are helped to view writing as an inherently difficult but nonetheless potentially manageable process. This is not to say that students never feel defensive or concerned when their work is reviewed in class, as shown by comments that have sometimes accompanied presentations of manuscripts such as "now you can have a chance to tear it apart." Nevertheless, over time, the process of writing and receiving critique does become less stressful for students because it is practiced repeatedly in a predictable and reasonably safe environment.

The process of repeated critiques also gives students the clear message that it is not only tolerable but actually useful and necessary to obtain consultation from one's peers concerning their work, which is a lesson that one hopes will generalize to their future professional functioning. The repeated deadlines that are an integral part of the seminar also provide a consistent structure with deadlines, which can facilitate more timely completion of writing projects.

The group support that is provided by the seminar has been particularly valuable in helping students cope with the distress stimulated by editorial review of their manuscripts and have helped formulate strategies for revision and resubmission. Moreover, hearing about the fate of their's and their colleagues' manuscripts also has helped students become familiar with the manuscript review process and also provides a useful forum for discussing professional/ethical issues concerning authorship and manuscript review. Students gain useful experience in reviewing others' work in progress by identifying problems, as well as communicating criticism in an empathic, constructive way.

In addition to the consistently positive feedback from students, evidence for the success of the writer's workshop has been evident in the fact that eight senior-authored manuscripts that were developed in the seminar over the past 3 years have been published and/or accepted in a peer-reviewed journal, most often JPP. I would encourage others to develop similar training formats to encourage students and faculty in pediatric and clinical child psychology to sharpen their skills in preparing scientific articles for publication.

REFERENCES

- Achenbach, T. M. (1991). *Manual for the Child Behavior Check List 14-18 and 1991 Profile*. Burlington, VT: Department of Psychiatry, University of Vermont.
- American Psychological Association (1997). Journals in psychology. A resource listing for authors. Fifth edition. Washington, D.C.: Author.
- American Psychological Association (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.
- American Psychological Association Summary Report of Journal Operations 1996. (1997). *American Psychologist*, 52, 908-909.
- Becker, H. S. (1986). *Writing for social scientists. How to start and finish your thesis, book, or article*. Chicago: University of Chicago Press.
- Bem, D. J. (1987). Writing the empirical journal article. In M. P. Zanna & J. M. Darley (Eds.), *The compleat academic: A practical guide for the beginning social scientist* (pp. 171-204). Hillsdale, NJ: Lawrence Erlbaum.

- Breakwell, G. M., Hammond, S., & Fife-Shaw, C. (Eds.). (1995). *Research methods in psychology*. Thousand Oaks, CA: Sage.
- Campbell, D. J., & Stanley, J. L. (1963). *Experimental and quasi experimental designs for research*. Chicago: Rand McNally.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Creswell, J. W. (1998). *Research design*. Thousand Oaks, CA: Sage.
- Day, R. A. (1988). *How to write and publish a scientific paper* (4th ed.). New York: ONYX Press.
- Drotar, D. (1989). Psychological research in pediatric settings. *Journal of Pediatric Psychology*, 14, 69–74.
- Drotar, D. (1997a). Relating family functioning to the psychological adjustment of children with chronic health conditions: What have we learned? What do we need to know? *Journal of Pediatric Psychology*, 22, 149–165.
- Drotar, D. (1997b). Intervention research. Pushing back the frontiers of pediatric psychology. *Journal of Pediatric Psychology*, 22, 415–424.
- Drotar, D. (1998). Training students for careers in medical settings: a graduate program in pediatric psychology. *Professional Psychology: Research and Practice*, 29, 402–404.
- Drotar, D., Lemanek, K., LaGreca, A. M., & Kazak, A. (1995). Case reports in pediatric psychology: Uses and guidelines for authors and reviewers. *Journal of Pediatric Psychology*, 20, 549–565.
- Drotar, D., Sturm, L., Eckerle, D., & White, S. (1993). Pediatric psychologists' perceptions of their work settings. *Journal of Pediatric Psychology*, 18, 237–248.
- Easterbrook, P. J., Berlin, J. A., Gopalan, R., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, 2, 868–872.
- Fine, M. A., & Kurdek, L. A. (1994). Publishing articles from a single data set. Issues and recommendations. *Journal of Family Psychology*, 8, 371–379.
- Fiske, D. W., & Fogg, L. (1990). But the reviewers are making different criticisms of my paper? Diversity and unique in reviewer comments. *American Psychologist*, 40, 591–598.
- Gardner, M. J., & Altman, D. G. (1986). Confidence intervals rather than *P* values: Estimation rather than hypothesis testing. *British Medical Journal*, 292, 746–250.
- Harter, S. (1985). *Manual for the Self-Perception Profile for Children*. Denver, CO: University of Denver.
- Haynes, S. N., Lemsky, C., & Sexton-Radek, K. (1987). Why clinicians infrequently do research. *Professional Publications Research and Practice*, 18, 515–519.
- Henson, K. T. (1995). *The art of writing for publication*. Boston MA: Allyn & Bacon.
- Holden, E. W., & Black, M. M. (1996). Psychologists in medical schools. Professional issues for the future: How are rank and tenure associated with productivity and satisfaction? *Professional Psychology: Research and Practice*, 27, 407–414.
- Home, P. D. (1988). Techniques for ensuring that your next paper is quite unsuitable for publication. *Journal of the Royal College of Physicians of London*, 22, 48–50.
- Howard, V. A., & Barton, J. H. (1986). *Thinking on paper*. New York.
- Huth, E. J. (1986). Irresponsible authorship and wasteful publication. *Annals of Internal Medicine*, 104, 257–259.
- Kazdin, A. E. (1980). *Research design in clinical psychology*. New York: Harper & Row.
- Kazdin, A. E. (Ed.). (1992). *Methodological issues and strategies in clinical research*. Washington, DC: American Psychological Association.
- Keith-Spiegel, P., & Koocher, G. P. (1985). *Ethics in psychology. Professional standards and cases*. New York: Random House.
- Kellogg, R. T. (1994). *The psychology of writing*. New York: Oxford University Press.
- LaGreca, A. M. (1994a). *Feedback checklist for submitted manuscripts*. Miami, FL: University of Miami.
- LaGreca, A. M. (1994b). *Working document: JPP feedback for authors*. Miami, FL: University of Miami.
- LaGreca, A. M. (1998). Reflections and perspectives on pediatric psychology: Editor's valedictum. *Journal of Pediatric Psychology*, 22, 759–770.
- Maher, B. A. (1978). A reader's writer's and reviewers guide to assessing research reports in clinical psychology. *Journal of Consulting and Clinical Psychology*, 46, 835–838.
- Nagata, D. U., & Trierweiler, S. J. (1996). Revising a research manuscript. In F. T. L. Leong & J. T. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants* (pp. 291–301). Thousand Oaks, CA: Sage.
- Norcross, J. C., Karg, R. S., & Prochaska, J. O. (1997). Clinical psychologists in the 1990s: II. *Clinical Psychologist*, 50, 4–11.
- Northridge, M. E., & Susser, M. (1994). Annotation: Seven fatal flaws in submitted manuscripts. *American Journal of Public Health*, 84, 718–719.

- Osipow, S. H. (1996). Dealing with journal editors and reviewers. In F. T. L. Leong & J. T. Austin (Eds.), *The psychology research handbook. A guide for graduate students and research assistants* (pp. 302–308). Thousand Oaks, CA: Sage.
- Parry, G. (1989). Writing a research report. In G. Parry & N. Watts (Eds.), *Behavioral and mental health research: A handbook of skills and methods* (pp. 105–122). Hillsdale, NJ: Lawrence Erlbaum.
- Peters, D., & Ceci, S. J. (1985). Peer review: Beauty is in the eye of the beholder. *Behavioral and Brain Sciences*, 18, 747–750.
- Peterson, C. (1996). Writing rough drafts. In F.T. Leong & J. Austin (Eds.), *The psychology research handbook. A guide for research assistants* (pp. 282–290). Thousand Oaks, CA: Sage.
- Range, L. M., & Cotton, C. R. (1995). Reports of assent and permission in research with children: Illustrations and suggestions. *Ethics and Behavior*, 5, 49–66.
- Roberts, M. C. (1992). Vale dictum: An editorial view of the field of pediatric psychology and its journal. *Journal of Pediatric Psychology*, 17, 785–805.
- Roberts, M. C., Lyman, R. D., Breiner, J., & Royal G. P. (Eds.). (1982). *Publishing child-oriented articles in psychology. A compendium of publication outlets*. Washington, DC: University Press of America.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.
- Rosnow, R. L., & Rosenthal, R. C. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, 44, 1276–1284.
- Squires, B. P. (1989). Descriptive studies: What editors want from authors and peer reviewers. *Canadian Medical Association Journal*, 141, 179–880.
- Squires, B. P. (1990). Statistics in biomedical manuscripts: What editors want from authors and peer reviewers. *Canadian Medical Association Journal*, 142, 213–214.
- Sternberg, R. J. (1988). *The psychologist's companion: A guide to scientific writing for students and researchers*. New York: Cambridge University Press.
- Sternberg, R. J. (1991). Editorial. *Psychological Bulletin*, 109, 34.
- Sternberg, R. J., & Gordeeva, T. (1996). The anatomy of impact. What makes an article influential? *Psychological Science*, 7, 69–75.
- Taylor, S. E. (1987). The progress and prospects of health psychology. *Health Psychology*, 6, 73–87.
- Thompson, R. J., Jr., & Gustafson, K. E. (1996). *Adaptation to childhood chronic illness*. Washington, DC: American Psychological Association.
- Wallgren, A., Wallgren, B., Persson, R., Jorner, V., & Haaland, F. A. (1996). *Graphing statistics and data. Creating better charts*. Thousand Oaks, CA: Sage.

17

Writing Research Reviews

DENNIS DROTAR

Many researchers in pediatric and clinical child psychology publish critical reviews of published research. The purpose of a research review is to define and clarify a scientific problem or issue; summarize previous investigations in order to inform the readers of the state of the art of research in an area of science; identify relative contradictions, gaps, and inconsistencies in the literature; and suggest an agenda to solve critical research problems [American Psychological Association (APA), 1994, p.5]. Such reviews serve several valuable functions: First and foremost, by summarizing what is known and not known about a field of scientific inquiry, a review identifies and clarifies salient research questions and sets directions for research that can advance the field. Because scientific progress builds on previous research, the ability to distinguish between what is already known versus what needs to be documented in a given area of research is a critical skill for researchers in clinical child and pediatric psychology. Researchers who ignore others' work not only run the risk of repeating methodological errors but of "re-discovering" existing knowledge.

Review articles also serve an important function for other researchers by providing a balanced critique of available work, underscoring methodological strengths and weaknesses of research, and suggesting new directions for research. Consequently, reviews can help to lay the foundation for new developments in research and theory by defining and giving direction to a relatively new field of inquiry and/or by taking critical stock of a field of science that is heavily traveled but needs to go in a different direction.

Researchers who learn to review and summarize research in an efficient, competent manner also gain an advantage in writing grants to secure resources for their research. For example, the background and significance section, which is a critical component of a grant proposal to the National Institutes of Health (NIH), requires researchers to critically evaluate the significance of research in a given

DENNIS DROTAR • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

field. The ability to critically evaluate relevant research, as well as identify the central problems and critical next steps in their field of interest, goes a long way to convince reviewers of an applicant's ability as a scientist (see Chapter 10, this volume). Similarly, foundations that fund research also are interested in a critical review of what has already been done in a given field of inquiry in order to determine what is new and significant about an applicant's proposed research (Carlson, 1995).

Most graduate students in pediatric and clinical child psychology are trained to review scientific literature and to critique research methods. However, the ability to critique, organize, and synthesize research in given field into a coherent review article involves special skills that are difficult to master and requires a great deal of practice to gain competency. While graduate students do obtain some experience in reviewing the literature in preparing their master's and dissertation research, such reviews tend to have a narrower scope than review articles and are generally not written for publication. Moreover, specialized technical skills and knowledge are needed to write different types of reviews, for example, meta-analyses (Cooper, 1982, 1984).

For the above reasons, there is a need for additional guidance and guidelines to help prospective authors to learn how to prepare review articles. Based on my experience as an author and interested consumer of review articles in pediatric and clinical child psychology, this chapter summarizes perspectives on how to synthesize information from the scientific literature and organize it into a coherent review article. In addition, using examples of descriptive reviews of research concerning topics in pediatric and clinical child psychology, pragmatic advice is given on preparing review articles. Readers who are interested in additional sources concerning the specialized technical and methodological issues involved in writing reviews including meta-analyses should consult other sources (Bem, 1995; Cooper, 1982, 1984, 1998; Cooper & Hedges, 1994; Glass, McGaw & Smith, 1981; Light & Pillemer, 1984; Rosenthal, 1995a,b). See also Chapter 18, this volume.

THE TASKS INVOLVED IN WRITING A REVIEW ARTICLE

Prospective authors of review articles face a daunting task. Reviewing the work of others requires a great many decisions. To help authors manage the preparation of a review article, the present chapter is organized around the tasks that need to be considered in preparing a review article for publication. Moreover, a list of helpful hints for writing a review article are summarized in Table 1.

DECIDING TO WRITE A REVIEW ARTICLE

The preparation of a review article involves a great deal of work. Consequently, before deciding to work on such a time- and energy-consuming project, prospective authors should consider whether their personal interest in the topic is sufficiently great to sustain a review and whether a review article is necessary to advance the field in their area of interest.

Table 1. Helpful Hints in Preparing a Review Article

General considerations and questions
Making the decision to write the review
Am I sufficiently interested in the topic to do the work?
Is a review article necessary to advance the field in the topic of choice?
Is there sufficient quantity and quality of research on the topic of choice?
What type of review article should I write?
Empirical
Theoretical
Methodological
Combination of the above
What is the topic of my review?
What content areas will I cover?
What key questions will I address?
Where should I publish my review article?
Specific suggestions in preparing the sections of a review article
Introduction
Decide on a clear focus of the review that is described in the introduction. Answer the question:
What is the purpose of the review?
Describe the significance of the topic: Answer the question: Why is this review necessary?
Describe the importance and unique or special contribution of the review. Answer the questions:
What new knowledge will be contributed by the review? Why is this knowledge important?
Define key terms and variables that will be used in the review.
Specify primary questions and the hypotheses that the review will address.
Consider relevant theory.
Method
Define clear boundary conditions for the review, e.g., content areas to be considered and dates of the review.
Define any specific criteria, e.g., methodological quality that were used to select studies for the review.
Describe procedures that are used to identify and locate studies that are reviewed.
Databases
Reference lists
Describe procedures that were utilized to summarize information in the review.
Presenting information concerning studies
Organize the review around tables that summarize data from the review.
Use themes to organize and present findings such as the following:
Content
Theoretical issues
Methodological problems
Utilize summaries of studies in preparing the review (see Table 2)
Keep track of and enter references and bibliography as you go along
Implications
Describe relevant implications:
Research
Methodological
Theoretical
Clinical
Develop new information that extends available research such as:
Novel integrations of theory
Suggestions for new research
Development of new methods

Personal Interest in the Topic of a Review

The motivation for the decision to prepare research review generally stems from a prospective author's strong personal interest in the topic. Most commonly but not always an author of a review article will be conducting or preparing to conduct research in the area of the work that is reviewed. For example, my motivation for the reviews that I have written on childhood chronic illness (Drotar, 1981, 1994, 1997) or failure to thrive (Drotar, 1989, 1995) stemmed primarily from my efforts to make sense of a disparate body of work that was scattered across different journals and also to provide direction to my own and others' work.

A strong personal interest in the topic of a review has both advantages and disadvantages that authors should recognize. Personal interest in a given area of research generally enhances both the liveliness of the writing and the utility of the work. On the other hand, one potential disadvantage of personal interest is the difficulty of maintaining objectivity, both in evaluating one's own work and others' research and points of view. Because of the potential for bias, authors who write research reviews on topics in which they have strong personal interest might wish to rely closely on colleagues for critical readings of their reviews.

Is a Review Article Needed to Advance the Field?

One of the most difficult questions for authors is whether it is advisable or necessary to publish a review. Making the decision about whether a review article is necessary to advance research in a topic area can be difficult, especially for reviews of research in relatively new or undeveloped areas. Consequently, before making the decision to write a review article for publication, prospective authors should carefully consider the quantity and quality of available research on a given topic. Some areas of research activity simply do not have a sufficient body of published work and/or research of a threshold level of methodological quality to warrant a review. In some areas, the quality of the work may be so limited that a review may not contribute much new information other than documenting obvious methodological problems. In such cases, the author may better spend his or her time doing the research that needs to be done rather than preparing a noncontributory review. Prospective authors who are uncertain about whether to prepare a review article may wish to contact colleagues who conduct research on the topic they wish to review to solicit their opinion about whether a review article is needed.

Prospective authors who are interested in populations and/or topics that have an extensive research literature face a different but equally difficult problem, that is, identifying a unique slant or niche for their review. See Barkley (1997) for an example of an excellent synthesis of research and theory on a clinical problem, attention deficit hyperactivity disorder, which has been extensively researched.

TYPES OF REVIEW ARTICLES

To maximize their efficiency, prospective authors of review articles should decide in advance about the type of review article they want to write. Review articles are heterogeneous, but they can be classified into several basic types. One

of these is an empirically based review that involves a detailed summary of the information or data contained in studies and a critique of the methods employed. Empirically based reviews include (1) traditional or content-focused reviewers that involve a narrative integration of research (see Drotar, 1997), and (2) meta-analysis, which is a quantitative method of organizing data. See Lavigne and Faier-Routman (1992) or Weisz, Weiss, Blicke, and Klotz (1987) for examples.

A second type is a theoretically based review, which is a critical integration of theory concerning a particular psychological issue that is intended to guide future research and/or practice (see Dix, 1991). A third type of review is methodological, which involves a critique of methods (see Meyer, 1997). Finally, a hybrid review employs some combination of the above (see Compas, 1987). In order to determine what type of review article is best for their purpose, prospective authors might wish to look at published reviews in different journals to familiarize themselves with alternative styles and formats.

What Type of Review Is Best?

What type of review should an author conduct? The answer to this difficult question depends on the author's purpose in writing the review, the questions that are addressed by the review, the level of development of a particular field of research, and the author's interests and skills. In making this decision, prospective authors should recognize that content-focused reviews, in which results of studies are presented and strengths and weaknesses discussed have significant limitations, especially for accurate syntheses of research on particular topics, such as treatment outcomes (Light & Pillemer, 1984; Rosenthal, 1995a,b). For example, many content-focused reviews lack adequate decision rules for evaluating treatment outcomes. In the absence of clear rules concerning how to evaluate or combine effects of different studies, content-focused reviews may reach faulty conclusions (Light & Pillemer, 1984). This is a significant problem with so-called "vote count" methods that tabulate numbers of studies with statistically significant versus nonsignificant findings (Cooper & Hedges, 1994) in order to reach a conclusion about a specific question, such as, "is behavior therapy with children effective?" For these reasons, meta-analyses have an important role in research reviews, most especially when the author's purpose is to review outcomes of treatment (see Durlak & Lipsey, 1991; Chapter 18, this volume).

Despite the above advantages, it are not always possible to conduct meta-analyses, especially in new areas of research where extensive data is not available. Moreover, authors who wish to synthesize and integrate research information in order to suggest new directions for research or draw out clinical implications may choose to organize their review around central themes and a theoretical perspective. For examples, see Dix (1991) for an interesting description of a model of parenting and Barkley (1997) for a theoretical perspective on attention deficit hyperactivity disorder.

WHAT SHOULD THE REVIEW FOCUS ON?

Identifying a specific focus is always important for an effective review but is especially critical if there have been prior reviews in a topic area and/or the area

has been extensively researched. In some areas of research, authors have many possible choices for review topics. For example, a review of research on conduct disorders in children and adolescents might focus on etiology, diagnosis, assessment, treatment, developmental issues, or an integrated synthesis of each of these areas (Bem, 1987). An author's decision concerning the specific focus of his or her review article should be influenced by significance or uniqueness of the topic, questions that are posed, and the value of the take-home message of the review (Sternberg, 1991). In addition to selecting the content area or domain of the review, a prospective author should describe the key questions that will be addressed by the article. For example, within the broad topic of assessment of conduct disorders in children, an author might want to consider the dilemmas that are involved in assessing conduct disorders in children, the types of clinical decisions that are made, alternative approaches to assessment of conduct disorders, studies of the reliability and validity of such assessments, and so forth.

One of the facts of life in preparing reviews is that it is usually difficult for prospective authors to clarify the specific questions to be considered in a review until they are relatively well along in the process of reviewing research. For this reason, in the early phases of their review, authors should anticipate that they may begin with several related topics and questions that will become more refined and tightly specified as the review proceeds.

WHERE SHOULD A REVIEW ARTICLE BE PUBLISHED?

Researchers who work with pediatric and clinical child populations have several options concerning publication outlets for review articles. The two major journals that publish research-based articles in pediatric and clinical child psychology, *Journal of Pediatric Psychology* (JPP) and *Journal of Clinical Child Psychology* (JCCP), place a clear priority on publishing original research papers rather than research reviews but do publish research review articles. Review articles are a regular feature in special sections of some journals such as the *Journal of Developmental Behavioral Pediatrics* or the *Journal of Child Psychology and Psychiatry*. Other journals that publish reviews and syntheses of research are *Behavioral Therapy*, *Psychological Bulletin*, *Psychological Review*, *Clinical Psychology Review*, *Clinical Psychology, Science, and Practice*, and most recently (March 1998) *Clinical Child and Family Psychology Review*. Relatively recent examples of reviews from JPP and JCCP include Spieth and Harris's (1996) review on assessing health-related quality of life in children and adolescents, Budd and Holdsworth's (1996) review of clinical assessment of minimal parenting competence, and Drotar's (1997) review of the relationship of parent and family functioning to the psychological adjustment of children with chronic health conditions.

THE STRUCTURE OF A REVIEW ARTICLE

Despite variability in purpose and authors' styles, reviews generally include five components: (1) an introduction and statement of the question that is to be addressed in the review; (2) description of methods; (3) presentation of the "data"

or summaries of information from the studies reviewed; (4) critique of findings/methods and/or theory reviewed; and (5) implications and recommendations for theory, research, or clinical management.

Introduction

The introduction to a review article should include a rationale for the importance of the review. State the goal of the review and the questions that will be addressed including key terms and definitions.

Provide a Rationale for the Importance of a Review

Authors of review articles need to consider four basic questions in their introductions: (1) Why is this review important or necessary? (2) What is the goal or purpose of the review? (3) What new knowledge will be gained by the review? (4) Why is the information contributed by the review important to the development of the field?

As is the case for any journal article (see Chapter 16, this volume), authors should clearly establish the need and rationale for their reviews in their introductions and not assume that their readers share their concept of the need for the review and significance of the topic. If previous reviews have been conducted in similar topics, authors should carefully describe the major findings and conclusions of previous work in order to highlight the needs in the field, set the stage for the specific questions that their review will address, and highlight the specific contribution that will be made (Light & Pillemer, 1984).

For example, to establish the need for a review of research concerning the impact of family functioning on adherence to treatment in childhood chronic illness, an author might first establish the importance of treatment adherence as a public health problem by documenting how noncompliance affects children's health and increases costs of medical care related to hospitalization and illness complications. The author also should describe the prevalence of compliance problems. Moreover, data concerning the relationship between family functioning and adherence to treatment should be cited. The rationale for the review also would be strengthened by research that links quality of family functioning to the level of children's adherence to treatment for a range of chronic health conditions.

The specific new knowledge that will be gained by conducting the review should be specified, for example, identification of key methodological issues in assessing adherence, suggestions for strategies of improving the design of studies that link family functioning and adherence, setting an agenda for future research, and so on. Readers also should be informed about why this knowledge is important, for example, improving the quality of research concerning adherence leading to more effective interventions.

State the Purpose of a Review and Highlight Critical Questions To Be Addressed

Developing a well-focused topic or theme for the review also will help to organize the author's task and clarify the purpose of the review for readers. Reviews

are much more effective if they focus on a specific topic area and set of questions, rather than a very broad topic area. Because reviews that address similar content areas can focus on very different questions, explicit statements of purpose are important. For example, one purpose of a review of child psychotherapy research would be to evaluate factors that influence treatment effects. An alternative, equally useful, but quite different focus within the same topic area would be to critically evaluate methodological approaches to evaluating treatment efficacy.

Authors also should specify the key scientific questions that their reviews will address. For example, if the purpose of a review is to assess the effects of intervention for children who are at risk for mental retardation, authors need to clarify such questions as which populations will be studied and what interventions will be considered? Alternatively, in a review of individual differences in response to intervention in the same population, authors might wish to consider such questions as under what conditions and with what samples are interventions particularly effective?

In order to articulate the central questions that will be addressed by their reviews most effectively, authors should clearly define the terms, definitions, and central variables that are essential to the primary questions posed by their reviews, and use these consistently throughout (Bem, 1987; Light & Pillemer, 1984). For example, for a review of research on early interventions with children at risk, one needs to define "early" (e.g., does this mean in utero, the first year of life, preschool? etc.), "intervention" (e.g., medical, psychological, social), and "at risk" for what? (e.g., child abuse, health problems, future cognitive, emotional problems, etc.).

The precision with which the reviewer frames and articulates their primary questions will help to identify the boundaries of the review and consequently the studies that are selected to be reviewed. Consider the example of the review of research on early intervention with children at risk referred to above. If the author defines "early" as occurring during the first year of life, "intervention" as counseling or guidance provided to parents, and the target risk outcome as children's emotional development, he or she can begin to set operational criteria for the studies that will be included.

Describe the Methods Used in the Review

To enhance replicability and clarity of their procedures, review authors should summarize the key dimensions of their methods including: (1) specific procedures to access studies that were reviewed (e.g., which specific databases were used to access articles that were selected to be considered in the review); (2) methods of selecting studies including criteria that were used to select studies together with the rationale for these criteria; and (3) analytic methods that were used to summarize information for the review (e.g., tabulating numbers of studies with certain characteristics and/or formal statistical methods, such as meta-analyses).

How Were Studies Accessed?

Authors of reviews should describe the way that they accessed studies in sufficient detail to allow others to replicate their methods. Researchers generally use computerized databases (e.g., Psych Lit or Medline to locate and select studies

to be included in reviews, (see Cooper, 1998, for a detailed description of databases utilized in reviews). However, it is important to recognize that the yield of these databases is highly dependent on the key words that are selected by the authors of articles in the database. If potentially applicable key words were not chosen by the author, the article will not be accessed. For this and other reasons, computerized searches are not a reliable way to identify all the relevant research literature in a content area, especially if the research topic is very broad, for example, psychotherapy or chronic illness in children or when research that is reviewed crosses disciplinary lines, as is often the case in pediatric and clinical child research. Durlak, and Wells (1997) noted that only one of every three entries in their computer-generated lists of studies of interventions were actually relevant to the study topic. Even more important, about two thirds of the relevant studies were not picked up by the author's computer search.

Given the problems associated with computerized searches, multiple methods should be used to search for relevant articles, including manual searches of journals that publish research pertinent to the review topic, examination of reference lists identified from published reviews and studies, consultation from individual researchers working in a field, and manual searches of multiple abstracting and indexing databases (Durlak & Lipsey, 1991).

How Were Studies Selected To Be Included In A Research Review?

In general, a review has two major kinds of boundary conditions: (1) the content area; and (2) inclusionary and exclusionary criteria for studies that are reviewed. The content area and specific criteria for selecting articles to be reviewed should be determined by the purpose and questions of the review as well as by the level of scientific development of research in a given topic area. For example, when reviewing studies in a fledgling area of research, it may be desirable to cast the review net broadly and set few criteria in order to clarify the salient research issues and methodological problems. On the other hand, in areas of research where scores of studies have been conducted, it is generally more desirable to narrow the field of the review by using more restrictive selection criteria.

Authors need to carefully set criteria for studies to be included and excluded in their research reviews. Decisions concerning criteria can involve difficult costs versus benefits considerations. For example, if reviewers do not set specific criteria to focus and limit the scope of their reviews, they are obliged to review each and every study that fits the review's content area. On the other hand, to the extent that authors exclude studies, they may create a biased selection of content (Light & Pillemer, 1984). Nevertheless, certain types of bias may be desirable, such as excluding studies that do not meet minimal methodological criteria (e.g., including control groups in studies of treatment outcome).

Consider Options for Inclusionary/Exclusionary Criteria in Selecting Studies

Light and Pillemer (1984) describe the options that are available to investigators to select studies for their reviews: These options are: (1) every available study on the topic, irrespective of publication status; (2) only published research; (3) using a panel of experts to select studies; and (4) selecting by study characteris-

tics, for example, types of control groups utilized. The latter two options could apply to published or unpublished studies, or both. Each of the above options for selecting studies has costs and benefits, which need to be considered in light of the specific purpose or questions considered in the review.

Review authors should consider the impact of selection biases concerning published studies (Greenwald, 1975; Smith, 1980) (see subsequent section on methodological limitations). Irrespective of methodological quality, published studies will generally show larger effects than unpublished studies. On the other hand, while inclusion of all available studies will limit the bias that is inherent in choosing only published studies, such a strategy is often highly impractical, given the problems of reliably locating unpublished studies. Moreover, the methodological quality of unpublished studies can be questioned because they have not had peer review.

Based on the above considerations, it is often desirable to set some criteria for the quality of methods in selecting studies to be included in a review. Unfortunately, the fact that there are several important dimensions of methodological quality, for example, design and sampling, measures, makes it difficult to operationalize ratings of methodological quality unless very basic criteria, such as, including only controlled studies, are used. Some authors (Glass et al., 1981) have suggested that for some purposes, liberal methodological criteria may be preferable in selecting studies because this strategy allows reviewers to address relationships between study methods and findings (e.g., do controlled studies of early intervention for economically disadvantaged children show a different pattern or strength of findings than uncontrolled studies).

Reviewers also may choose to specify key methodological categories as inclusion criteria. For example, in a primary review of the psychological adjustment of children with chronic physical illness, an author might select studies on the basis of study design (e.g., studies that employed physically healthy controls or those that used objective, well-standardized measures of psychological adjustment). Irrespective of the specific criteria that they use to select studies to be reviewed, authors should give clear rationale for the criteria and operational definitions in order to enhance replication of methods by other researchers.

The time frame reviewed studies is an important but difficult criterion for several reasons: Given the relatively recent availability of computerized search procedures, it is impractical to begin a review with the first published study that was ever conducted on a topic. It is also difficult to identify an objective marker to determine when the review should begin. Some authors may find it useful to begin their review at the date of the last major review of the research topic in question, that is, if one had been conducted. Others may choose a marker date to begin their review based on data that there were few relevant and/or high-quality studies conducted prior to that date.

Specify Primary Hypotheses

Because they provide clarity and structure to the presentation of information, specification of hypotheses are critical in meta-analytic reviews but should be carefully considered in any review. Specific hypotheses also limit the problem of overcapitalizing on chance findings that can characterize more exploratory-driven

reviews (Light & Pillemer, 1984) and enable authors to highlight the specific scientific contribution that will be made by their reviews.

Authors should carefully choose their hypotheses to facilitate maximal impact of their reviews. For example, because psychotherapy with children has received considerable empirical scrutiny in several reviews (Weisz et al., 1987), a review that focuses on evaluating the hypothesis, "child psychotherapy is effective" would not be contributory. In order to make a scientific contribution in this field of research, a prospective review author would need to consider a more specific hypothesis such as "are the effects of psychotherapy moderated by key variables, for example, quality of family functioning?"

Articulate Relevant Theory

In developing their reviews, authors should carefully consider and articulate theoretical issues that are relevant to the topic of their research. A detailed consideration of theory will help integrate the research that is reviewed and present it coherently. Moreover, close attention to theoretical issues that relate to research that is reviewed will add to the value of the review by developing theory, which is an important need in pediatric and clinical child psychology research (Wallander, 1992).

Presenting the Findings of a Review

Organizing Findings

The organization and presentation of the information from studies that are reviewed present a formidable challenge for authors of review articles. Several formats for presentation of findings are available. One common choice is to group studies under specific headings that are judged to best represent the content areas of the review. For example, it may be effective to organize a review around topics that reflect the major themes that have emerged in the review, salient methodological issues, or theoretical issues (see Dix, 1990, for an excellent example of this latter approach). Because the most effective method for presenting and organizing the information initially may not be apparent, authors may wish to experiment with several formats of presenting information before deciding on the most effective one (Cooper, 1998).

Using Tables to Summarize Information

To help readers understand and process information about study characteristics and findings, it is desirable to use a table to summarize relevant information for example, authors, sample size, design, measure, findings, and so on, from the studies that are reviewed. Given the large numbers and complexity of studies that are often included in reviews, the choice of a table format that conveys the necessary information yet does not overwhelm the reader can be difficult. For this reason, it may be helpful to ask colleagues who are not familiar with the content area of the review to look at the tables from the standpoint of utility and readability. See reviews by Lavigne and Faier-Routman (1992) and Weisz, et al. (1987) for examples

of useful tables and Cooper (1998) for an informative discussion of data presentation in review articles.

Develop a Critique of Findings

To be most effective, the presentation of findings from studies that are reviewed should be integrated with the critique of these studies (see previous section). In presenting information concerning their critique of studies, authors face difficult decisions about how much and what to say about individual studies. Information from a group of studies that reach similar conclusions can be quickly summarized in several statements. Other studies are best presented in more depth as exemplars of special methodological and/or theoretical contributions. Presentation of patterns of findings that emerge from individual studies or a description of central themes that characterize a body of literature are clearly preferable to a study-by-study description, which is difficult for readers to integrate and is not terribly exciting.

Describe Implications and Recommendations for Research, Theory and Practice

The author's specific implications and recommendations are critical to the success of a review because they are the primary "take-home" messages to readers. Nevertheless, the absence of useful take-home messages are a common problem in reviews as well as in original research articles (Sternberg, 1991). The different types of implications and recommendations that may be drawn from a review article—methodological, research, theoretical, and clinical—are now described.

Methodological Implications

One of the reviewer's primary tasks, especially in reviews with a primary methodological focus, is to identify the design and methodological problems that pose the most central threats to the validity of study findings (Campbell & Stanley, 1963), to alert readers to these problems, and to outline an agenda for research that will advance the field by addressing these problems.

Research Implications and Recommendations

The most useful reviews also contain detailed recommendations for strategies to manage or overcome recurrent methodological problems that characterize a given area of research. In order to give concrete illustrations of relevant points, which many readers find to be helpful, it is often very useful to highlight examples of studies that in the author's view have addressed important methodological problems in helpful ways. The most useful reviews also describe directions for research that in the reviews opinion, are most likely to advance scientific knowledge in the area of work that is reviewed. Reviewers have an important opportunity to shape research in their field of interest by underscoring new research directions, by highlighting examples of cutting edge research, stating specific hypotheses that should be tested, and/or new populations that are important to study.

Theoretical Implications

Reviewers can make a valuable contribution by recommending the theories that are needed to advance scientific knowledge by using research to develop theory and by developing creative application of theory to a new area of research. An important function of review articles is to integrate seemingly multiple theoretical perspectives into a coherent framework that serves to unify the field (Sternberg, 1991). See Barkley (1997) for an example of such integration.

Clinical Implications

While every research review need not have clinical implications, wherever possible authors should describe the potential clinical implications of the research they review. Authors who have extensively reviewed research in a field have an excellent opportunity to carefully consider the potential clinical implications (e.g., for assessment and intervention) of the work they have reviewed, to integrate research findings with important clinical problems and issues, and/or to suggest ways that research can be made more clinically relevant. Kazdin's (1988) and Weisz et al.'s (1987) reviews of psychotherapy research with children are excellent examples of review articles that have made important recommendations to the clinical relevance of research. See also Drotar (1989) for a more specific example of clinical implications based on a review of behavioral assessment and diagnosis in failure to thrive, or Squires, Nickel and Eisert (1996) for clinical implications based on a review of screening for early detection of developmental problems.

COMMON DILEMMAS IN WRITING REVIEW ARTICLES

Articulating the Novel Contribution of a Review Article

One of the dangers of writing a review paper is that authors can become so thoroughly immersed in published literature that they fail to develop a "point of view" about their topic.

The best reviews are not mere recapitulations of existing work but serve to help researchers see an area of research in a new light. Consequently, authors of reviews should push themselves to take some chances in articulating their own critique of research and suggestions for future research directions (Sternberg, 1991). Authors of review articles can extend scientific knowledge in several different ways: by using the research findings to develop and extend theory in novel ways, by integrating research and theory in fields that have been previously seen as separate, or by making novel suggestions for research directions or methods to solve recurrent methodological problems that have been identified.

Appreciating the Evolution of Scientific Work

Reviewers are in the position of being "Monday morning quarterbacks," who have the considerable advantage of hindsight. While such hindsight can be instructive, it also can lead review authors into various traps that can limit the effectiveness of their reviews. One trap is the failure to appreciate the evolution of

the state of the art of science in the area of research that is reviewed. Science builds progressively on previous methods and theoretical perspectives. Consequently, work that was previously considered to be state-of-the-art research is almost always disadvantaged compared with more recent work. Consider the following example in the field of pediatric psychology: Research that studied the psychological adjustment, such as, self-esteem of children with a particular chronic health condition, for example, children with diabetes who were studied at a single site was novel in the 1970s early in the field of pediatric psychology when the adjustment of children with chronic health conditions was not well documented. Now that there have been a great many studies of psychological adjustment of children with chronic health conditions (Lavigne & Faier-Routman, 1992), the field has advanced to the point that such studies are no longer useful unless they study a new outcome, process, or design (e.g., prospective methods) or a new sampling strategy (e.g., multisite research) (Drotar, 1994). In order to give readers their balanced perspective, reviewers need to understand the evolution of a field of research and appreciate the context, (e.g., time period of studies reviewed) and methods that were employed at these times.

Reviewing Critically yet Constructively

A related problem in writing a review article is the tendency to focus on the weaknesses of studies that are reviewed without appreciating their strengths. In their zeal to provide a thorough critique, reviewers may not sufficiently recognize examples of excellent research. Recognition of positive examples of research is valuable in several respects. First, it underscores positive models of research that other researchers can strive to emulate. Second, it provides encouragement to researchers and serves to stimulate new research. Third, such recognition allows for a balanced presentation of findings that can be more accurate and interesting than a cataloging of faults and problems.

Focusing a scientific review solely on what was wrong with a given area of research may be technically correct but may not do sufficient justice to the inherent problems involved in some areas of research. Moreover, a nihilistic tone of a review may actually discourage researchers from taking on potentially important challenges in a difficult area of research. Finally, overly critical findings and conclusions may actually limit a review article's chance of being accepted for publication if journal editors and reviewers are persuaded that the area of research is too underdeveloped or problematic to warrant a published review.

Recording Information from Studies That Are Reviewed

One of the important issues that review authors need to consider is how best to record and synthesize information from the studies that they review. Because the sheer quantity of information that is contained from studies that are reviewed can be overwhelming, authors need a way to organize and retrieve this information. In writing reviews, I have found it useful to abstract the relevant information from a review article into a summary sheet or table that includes relevant information such as authors, journal, sample size and selection, demographic characteristics, relevant theory and hypotheses, methods (e.g., cross sectional, prospective), mea-

sures, method of data analysis, major findings, conclusions, strengths, and methodological problems. See Table 2 and Cooper (1982) for sample formats for summaries.

Structured formats such as those suggested above to summarize information that is reviewed serve several purposes: They provide a way to access information that is more efficient than referring back to the original articles each time. They also can help to focus attention on the most relevant characteristics of the articles reviewed. Finally, such formats also can be used to construct summary tables of the articles reviewed which can be very helpful to authors and eventually to readers.

Using Qualitative Data

Whether and how best to use qualitative information, for example, single case studies, expert judgment, or narrative reviews of research and theory (Light & Pillemer, 1984) in a review article is an interesting and controversial issue. Some researchers argue for the need to synthesize all available information in a given topic (Cook & Levitan, 1980). This strategy can be particularly important in emerging areas of research where the primary purpose of the review is to identify salient research topics and directions. Qualitative information, such as that obtained from some case studies, may help interpret new phenomena, which can be instructive in evaluating scientific information (Kazdin, 1980; Light & Pillemer, 1984). Because there are no hard and fast rules about including qualitative information in reviews, authors will need to use their judgment to determine what is best for their purpose. Nevertheless, authors who do include qualitative information in their reviews should provide a clear rationale for doing so.

METHODOLOGICAL LIMITATIONS OF RESEARCH REVIEWS

Reviewers and the consumers of reviews need to consider the inherent methodological limitations of reviews as a method of summarizing scientific literature (see Cooper & Hedges, 1994; Durlak & Lipsey, 1991). Although these problems are difficult to surmount, because they reflect practices in publication of articles, they need to be considered in interpreting the findings from reviewers.

Publication Bias

One of the most important and well-recognized limitations of reviews of published research relates to the bias against publishing studies that have nonsignificant findings. Many have argued convincingly that the field is being limited by an undue dependence on statistical significance (Rosenthal, 1995a). Nevertheless, the fact remains that journal reviewers and editors are more likely to publish studies that achieve statistical significance than those that did not, even if their study designs were equivalent or better than studies with nonsignificant findings (Easterbrook, Berlin, Gopalan, & Matthews, 1991).

A related problem in interpreting findings based on research reviews concerns the file drawer problem (Rosenthal, 1979), that is, the published literature is a biased underestimate of the numbers of studies that actually have been conducted on a particular topic. Researchers who did not obtain statistically significant

Table 2. Sample Format for Abstracting Information from Studies

Title:			
Authors:			
Journal:			
Population:	Type of group (e.g., chronic illness)		
	Sample size		
	Group 1 _____		
	Group 2 _____		
	Group 3 _____		
	Group 4 _____		
Design	<input type="checkbox"/> Single group <input type="checkbox"/> Cross-sectional <input type="checkbox"/> Multiple group <input type="checkbox"/> Prospective		
Measures/variables	Independent	Dependent	
	1. _____	1. _____	
	2. _____	2. _____	
	3. _____	3. _____	
	4. _____	4. _____	
	5. _____	5. _____	
	6. _____	6. _____	
	7. _____	7. _____	
	8. _____	8. _____	
Where specific hypotheses made?	<input type="checkbox"/> Yes <input type="checkbox"/> No		
If yes, list:	1. _____ _____ 2. _____ _____ 3. _____ _____		
Were hypotheses clearly derived from theory or conceptual framework?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Not sure		
If yes, what theory?	<hr/> <hr/> <hr/> <hr/>		
Sample selection procedures:	<hr/> <hr/> <hr/> <hr/>		
Were differences between participants and nonparticipants analyzed?	<input type="checkbox"/> Yes <input type="checkbox"/> No		
If yes, what differences were found?	<hr/> <hr/> <hr/> <hr/>		

Table 2. (Continued)

Demographic characteristics	
Race/ethnicity	_____
SES distribution	_____
Male/female	_____
Age/mean	_____
SD	_____
Range	_____
Method of analysis:	
ANOVA	_____
Regression	_____
Other	_____
Findings	
Significant independent variables	
1.	_____
2.	_____
3.	_____
4.	_____
5.	_____
6.	_____
7.	_____
8.	_____
Second analysis (if relevant)	
1.	_____
2.	_____
3.	_____
4.	_____
5.	_____
6.	_____
7.	_____
8.	_____
Conclusions	
1.	_____

2.	_____

findings may not choose to publish a study because they anticipate that the study will not be published. Alternatively, they may have tried to publish the paper but were unsuccessful and may not resubmit for publication. The net effect is that the published literature reflects only a subset of the studies that have been conducted on a particular topic.

Depending on the topic, the combination of publication bias and the file drawer problem could seriously limit and even distort the conclusions that can be drawn from research reviews. For example, consider the impact of these problems on a review of the efficacy of treatment. What if 100 studies were conducted on the effects of cognitive behavioral therapy on children with attention deficit hyperactivity disorder, but only 20 of these, the ones with the most positive treatment effects, were published. On the basis of the published studies, one would conclude

that the intervention had extremely positive effects. However, if all available data on treatment studies that had been conducted were reviewed (assuming a reasonable level of methodological quality in the unpublished reports), the reviewer's conclusions concerning the efficacy of treatment would be quite different.

The problems of inference associated with publication bias and the file drawer problem are very difficult to prevent, although it is possible to develop estimates of their effects (Rosenthal, 1995a). To help reduce the impact of these problems, reviewers might wish to contact researchers who they know are working in the area of research under consideration and invite their data to be reviewed. This proactive strategy has the advantage of providing a more comprehensive review of research than relying solely on published data. Nevertheless, even if a review author is well-connected to other researchers in a given area, he or she cannot be certain they can locate all relevant research that is being done on a particular topic. Moreover, some researchers who are contacted may choose not to share their data before they are published.

Generalizability Of Findings From Research Reviews

The generalizability of findings that are based on research reviews is another important issue. In evaluating the generalizability of data included in reviews, it is important to consider how the studies were selected for the review. Moreover, the characteristics of samples of studies that were included in a review should be carefully considered. For these reasons, a detailed description of research participants who are included in the studies selected for the review will provide important information about the potential generalizability of conclusions to other samples. For example, if studies of psychotherapy of children that were published and reviewed included only parents from low socioeconomic status, inferences to the general population of children would not be warranted. In this regard, the lack of inclusion of children and families from ethnic minorities in research poses a continuing threat to the generalizability of research findings in the fields of clinical child and pediatric psychology (see Chapter 8, this volume).

The reviewers' careful attention to the characteristics of participants in the studies that are reviewed also can lead to important conclusions and recommendations. For example, Kazdin's (1988) review of child psychotherapy research emphasized the narrow nature of treatment methods and clinical problems that have been evaluated and recommended broader, more clinically relevant treatment research.

TRAINING IN WRITING REVIEW ARTICLES

Out of inexperience, authors of review articles generally underestimate the work that is necessary to complete a review and the strategies that are useful in the preparing a review article. Consequently, formal training in the preparation of review articles is necessary. Our experience has suggested some training experiences to enhance students' skills in writing review articles. For example, the graduate training program at Case Western Reserve University changed the format of the special qualifying exams to include preparation of a publishable scholarly review of research under the tutelage of a mentor as a requirement for graduation.

By making a scholarly research review a formal requirement for graduation, the program recognizes the review and synthesis of research as an important skill and also builds in opportunities for students to practice skills in reviewing research and preparing a review under mentorship.

We have used a writer's workshop seminar format to help facilitate students' skills in developing research reviews (see Chapter 16, this volume). The seminar, which includes students and occasionally faculty who are preparing reviews for publication, provides a forum for discussing strategies to manage the difficult but predictable problems of developing a research review, for such, choosing a topic, deciding on criteria, how to present the information from student, drawing conclusions, and so forth. In each and every review article that has been prepared in this seminar, the level of time and energy required in preparing the research review has far exceeded the authors' expectations. Consequently, the availability of group support from members of the seminar has facilitated the quality of the manuscripts that were prepared. Thus far, two of the review articles prepared by students in this workshop over the past several years have been published (Ievers & Drotar, 1996; Nassau & Drotar, 1998). Others are in various phases of preparation and resubmission. Recently, faculty and trainees in the department of pediatrics who were preparing review articles also have participated in this seminar. I would encourage others to utilize similar training experiences to facilitate student and faculty preparation in writing review articles.

REFERENCES

- American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.
- Barkley, R. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65–94.
- Bem, D. J. (1987). Writing the empirical journal article. In M. P. Zanna & J. M. Darley (Eds.), *The complete academic. A practical guide for the beginning social scientist* (pp. 171–204). Hillsdale, NJ: Lawrence Erlbaum.
- Budd, K. S., & Holdsworth, H. J. (1996). Issues in clinical assessment of minimal parenting competence. *Journal of Clinical Child Psychology, 25*, 1–10.
- Campbell, D. J. & Stanley, J. L. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Carlson, M. (1995). *Winning grants step by step*. San Francisco: Jossey Bass Publishers.
- Compas, B. E. (1987). Coping with stress during childhood and adolescence. *Psychological Bulletin, 101*, 393–403.
- Cook, J. D. & Levitan, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta analysis. *Journal of Personality, 48*, 449–472.
- Cooper, H. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research, 52*, 291–302.
- Cooper, H. (1984). *The integrative research review: A social science approach*. Beverly Hills, CA: Sage.
- Cooper, H. (1998). *Synthesizing research. A guide for literature reviews*. Thousand Oaks, CA: Sage.
- Cooper, H., & Hedges, L. V. (1994). *Handbook of research synthesis*. New York: Russell Sage Foundation.
- Dix, T. (1991). The affective organization of parenting: Adaptive and maladaptive processes. *Psychological Bulletin, 110*, 3–25.
- Drotar, D. (1981). Psychological perspectives on chronic childhood illness. *Journal of Pediatric Psychology, 6*, 211–228.
- Drotar, D. (1989). Behavioral diagnosis in nonorganic failure to thrive: A critique and suggested approach to psychological assessment. *Journal of Developmental and Behavioral Pediatrics, 10*, 48–57.

- Drotar, D. (1994). Psychological research with pediatric conditions: If we specialize, can we generalize? *Journal of Pediatric Psychology*, 19, 403–414.
- Drotar, D. (1995). Failure to thrive (growth deficiency). In M.C. Roberts (Ed.) *Handbook of pediatric psychology* (pp. 516–535). New York: Guilford Press.
- Drotar, D. (1997). Relating family functioning to the psychological adjustment of children with chronic health conditions: What have we learned? What do we need to know? *Journal of Pediatric Psychology*, 22, 149–165.
- Durlak, J. A., & Lipsey, M. W. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology*, 19, 1291–332.
- Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A metanalytic review. *American Journal of Community Psychology*, 35, 115–152.
- Easterbrook, P. J., Berlin, J. A., Gopalan, R., & Matthews, D. R. (1991). Publication bias in clinical research. *Lancet*, 2, 868–872.
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Greenwald, N. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, 85, 845–847.
- Ievers, C., & Drotar, C. (1996). Family and parent functioning in cystic fibrosis. *Journal of Developmental and Behavioral Pediatrics*, 12, 48–55.
- Kazdin, A. E. (1980). Research design in clinical psychology. New York: Harbor & Row.
- Kazdin, A. E. (1988). Child psychotherapy. Developing and identifying effective treatments. New York: Pergamon Press.
- Lavigne, J. V. & Faier-Routman, J. (1992). Psychological adjustment to pediatric physical disorders: A meta-analytic review. *Journal of Pediatric Psychology*, 17, 133–158.
- Light, R. J., & Pillemer, D. B. (1984). *Summing up the science of reviewing research*. Cambridge, MA: Harvard University Press.
- Meyer, G. I. (1997). Assessing reliability: Critical collections for a critical examination of the Rorschach Comprehensive System. *Psychological Assessment*, 9, 480–489.
- Nassau, J. H., & Drotar, D. (1998). Social competence among children with central nervous system conditions: A review. *Journal of Pediatric Psychology*, 22, 771–793.
- Rosenthal, R. (1979). The “file drawer” problem and the tolerance for null results. *Psychological Bulletin*, 86, 638–641.
- Rosenthal, R. (1995a). *Meta-analytic procedures for social research* (rev. ed.). Edition, Newbury Park, CA: Sage.
- Rosenthal, R. (1995b). Writing meta-analytic reviews. *Psychological Bulletin*, 118, 183–192.
- Spieth, L. E., & Harris, C. V. (1996). Assessment of health related quality of life of children and adolescents an integrated review. *Journal of Pediatric Psychology*, 21, 175–193.
- Squires, J., Nickel, R. E. & Eisert, D. (1996). Early detection of developmental problems: Strategies for monitoring young children in the practice setting. *Journal of Developmental and Behavioral Pediatrics*, 17, 420–427.
- Sternberg, R. J. (1991). Editorial. *Psychological Bulletin*, 109, 3–4.
- Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation and Education*, 4, 22–24.
- Wallander, J. L. (1992). Theory-driven research in pediatric psychology: A little bit on why and how. *Journal of Pediatric Psychology*, 17, 521–536.
- Weisz, J. R., Weiss, B., Blicke, M. D., & Klotz, A. K. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, 55, 542–549.

18

How to Evaluate a Meta-Analysis

JOSEPH A. DURLAK

YOUR DILEMMA

You encounter a meta-analysis evaluating treatments for child and adolescent problems. The abstract suggests that the reviewed treatments are highly effective. The author states that current results challenge many widely held beliefs in the field and have strong implications for future research and practice. Your dilemma is that you know nothing about meta-analysis. How can you tell if this meta-analysis is trustworthy? This review could be very important but perhaps it was not conducted carefully enough to justify its findings and conclusions. What do you do?

This brief chapter cannot completely resolve your dilemma but it can help you understand the key evaluative criteria for judging the quality of a meta-analysis. As a result, after reading this chapter you should have an understanding of the basic features of a good meta-analysis and where to go for further details. In fact, before discussing evaluative criteria, several excellent resources are noted that explain the basics of meta-analysis (Light & Pillemer, 1984; Rosenthal, 1995; Wolf, 1986), discuss some of the choice points and judgment calls involved (Matt & Navarro, 1997; Nurius & Yeaton, 1987), provide details on technical and statistical aspects (Cooper & Hedges, 1994; Durlak & Lipsey, 1991; Hedges & Olkin, 1985; Hunter & Schmidt, 1990; Rosenthal, 1991), or walk the reader through a typical meta-analysis (Durlak, 1995).

WHAT IS META-ANALYSIS?

In brief, meta-analysis is a method of evaluating research findings in a quantitative fashion. Results from each study are transformed to a common metric—the effect size—and then the distribution of the effect size data is examined and analyzed. The two major types of effect sizes are r , or the product moment correlation, and the standardized mean difference, which goes by different designations such as d , $d+$, g , or ES (for effect size). This chapter focuses on reviews of outcome research comparing the status of treatment and control groups and discusses only standardized mean effect sizes (using ES as the abbreviation), which are the most commonly calculated metrics in such meta-analyses. Across the range of effects typically obtained in the psychological literature, values for r and ES can easily be transformed into the other metric: r is equal to one half the value of ES.

ESs are calculated so that higher values indicate that the experimental or intervention group outperforms or is superior to the control group. ESs also can be calculated when two treatments are being compared and in a pre–post single-group design, but these data should be analyzed separately. The former ESs typically are much smaller than treatment versus control group effects and the latter are usually much higher. ESs also apply to single-subject designs (see DuPaul & Eckert, 1997), but these ESs are not directly comparable to the ESs discussed here because the former are based on within- rather than between-subject data.

An ES from an individual study can be of any magnitude, but two thirds of all mean ESs obtained in reviews of education, mental health, and social science research have been between 0.19 and 0.75, and 83% of the mean ESs have been greater than 0.20 (Lipsey & Wilson, 1993).

Basically, the challenge for the meta-analyst is to explain the variability that usually occurs when ESs across studies are combined. Different studies produce different results. If studies consistently produced similar unequivocal outcomes, then there would be little need for an integrative synthesis. However, the outcomes and the characteristics of reviewed studies vary simultaneously, and the meta-analyst must make sense of these variations in as objective and unbiased manner as possible. Moreover, the possibility that various sources of error are responsible for the different outcomes achieved across studies also needs to be explored.

CRITERIA FOR EVALUATING A META-ANALYSIS

While there are several important aspects to meta-analysis, the focus here is on seven salient criteria that apply to most situations. Other evaluative schema for meta-analysis exist (see Cooper & Hedges, 1994; Light & Pillemer, 1984; Swanson, 1996). The seven criteria are listed in Table 1 in the form of evaluative questions and suggest the features that are present in higher-quality reviews. These criteria are not exhaustive and their presence does not guarantee a good review, but in general one can place greater confidence in meta-analysis that satisfies these criteria.

Are Specific Hypotheses Present?

The purpose of many early meta-analyses was primarily descriptive, that is, to summarize the results of a literature and to describe some of the characteristics of relevant studies. The rationale for these descriptive meta-analyses is now past

Table 1. Seven Criteria for Judging the Quality of a Meta-analysis^a

-
1. Are specific hypotheses present?
 2. Are representative studies evaluated?
 3. Are appropriate analyses conducted?
 4. Is the meaning of outcomes sufficiently explained?
 5. Is the practical significance of outcomes assessed?
 6. Are findings and conclusions appropriately qualified?
 7. Are specific directions for future research offered?
-

^aNOTE: Alternative schema for evaluating a meta-analysis are available in Cooper and Hedges (1994), Light and Pillemer (1984), and Swanson (1996).

because this information can be subsumed in more powerful so-called explanatory meta-analyses. Therefore, a high-quality meta-analysis seeks to explain the reasons for the obtained effects, that is, to identify potential moderators of outcome. To explain the results, it is important that the meta-analyst have a good prior understanding of the theoretical and empirical literature. This understanding will help formulate specific hypothesis that can be tested. All research areas have some inconsistent findings, some controversies, or some competing theories or explanations for the results that have been obtained to date. A good meta-analysis will attempt to resolve these issues to the extent possible.

A priori hypotheses are strongly preferred in a meta-analysis. Otherwise, many statistical tests can be conducted and confidence in the results diminishes because the more tests one conducts, the more likely it is that some will achieve statistical significance by chance.

Are Representative Studies Reviewed?

The findings from a meta-analysis are only as good as the data on which they are based, and these data should be drawn from as representative a set of studies as possible. It is up to the meta-analyst to convince the reader that studies were not obtained in a biased fashion. In doing so, two issues are prominent (for additional issues see Durlak & Lipsey, 1991; Tabak et al., 1991): (1) were multiple methods used to locate studies? and (2) were unpublished studies evaluated.

In most cases, representative studies are best obtained through multiple search strategies that involve computer searches, manual searches, and inspection of the reference lists of included studies and previous research reviews. As tedious as it sounds, an excellent way to locate studies is to go through multiple volumes of the journals that tend to publish articles in the target area and examine every study. Beware of meta-analyses that depend primarily on computer searches to locate relevant studies. Computer searches are notoriously unreliable because they typically yield many irrelevant studies and insufficient numbers of relevant studies (Durlak & Lipsey, 1991). In some cases, computer searches have identified as little as 3 to 6% of the relevant research (see Lösel, 1991).

Computer searches cannot be trusted for at least three reasons. First, computer databases do not include all publication outlets; second, the index terms used in such databases might not be specific enough for the literature domain in question;

and third, some subjectivity is involved when indexing each particular study. Two highly similar studies can be indexed differently, and thus both will not be retrieved in the same computer search.

For instance, a reviewer might want to search for school-based studies of elementary school children using cognitive-behavioral therapy to promote social skills. There is no guarantee, however, that most relevant studies will be retrieved in a computer search because the terms, *school-based*, *elementary school children*, *cognitive-behavioral*, and *social skills* and their variants are broad terms that will not be consistently captured in most indexing systems.

It is extremely helpful to evaluate a sample of unpublished studies that can be obtained using the same methods to locate published studies. Evaluating unpublished work is important because of publication bias that often characterizes research in medicine, education, and the social science, published studies tend to yield higher ESs than unpublished ones (Dickersin, 1997). Publication bias occurs because of the tendency on the part of editors and reviewers to reject for publication studies that yield nonsignificant results and to the reluctance on the part of researchers to submit all their studies for possible publication, in particular, those which do not obtain positive results.

Although publication bias is always a possibility, it has not been found in every meta-analysis (e.g., Durlak & Wells, 1997). As a result, the existence and degree of publication bias are empirical issues that should be examined in each review. Without a sample of unpublished studies in the meta-analysis; however, publication bias cannot be evaluated.

Unpublished dissertations probably represent the best source for unpublished work, because *Dissertation Abstracts* contains a listing of dissertations completed each year at United States and Canadian institutions (so the existence of unpublished literature is publicized and can be scrutinized); most dissertations can be obtained free of charge through interlibrary loan; and dissertations frequently contain more procedural and statistical details than published papers. One can try to obtain unpublished convention papers and file drawer studies (completed studies that are languishing in some researcher's files), but investigators do not readily cooperate with requests for copies of their unpublished work.

There are some striking examples in which reviews of ostensibly the same literature contain very different study samples. For example, two early meta-analyses of the child therapy literature, one with 64 outcome studies (Casey & Berman, 1985); and the other with 105 (Weisz, Weiss, Aliche, & Klotz, 1987), had only 24 studies in common. Six reviews of student evaluations of teachers' effectiveness included 3, 6, 14, 16, 18, and 41 studies, and no single study appeared in every review (Abrami, Cohen, & d'Apollonia, 1988). Similarly, no single study appeared in each of four reviews of school-based drug prevention programs (Hansen, 1992). The above findings make it difficult to determine which study sample in which review (indeed, if any at all) adequately represent the literature.

Are Appropriate Analyses Conducted?

Three important aspects of statistical analyses are highlighted here: (1) were studies weighted prior to analyses; (2) was one effect size from each study used for each research question; and (3) were rival explanations for outcomes evaluated

Weighting Effects

Hedges and Olkin (1985) have emphasized that ESs should be weighted by the inverse of their variance before any analyses are conducted (p. 111 provides the equation that is heavily influenced by sample size). This approach is now commonly used, although other weighting schemes are possible (Cooper & Hedges, 1994; Hunter & Schmidt, 1990).

Weighted mean effects are preferred because weighted effects provide a more accurate estimate of true population effects.* Therefore, it is important that the meta-analyst weight studies before conducting any statistical analyses. In three meta-analyses in which the authors made separate calculations, weighted mean effects were from 23 to 33% lower than unweighted effects (see Beelmann, Pfingsten, & Lösel, 1991; Nearpass, 1990; Weisz, Weiss, Han, Granger, & Morton, 1995). Based on this information, if the meta-analyst only reports unweighted means, it might be wise to reduce the obtained effect by 25% or so to gain a more conservative estimate of the magnitude of effect.

One Effect per Research Question

In many early meta-analyses, effect sizes from every outcome measure were included in the analyses (e.g., Prout & DeMartino, 1986; Smith & Glass, 1977). Because the number of outcome measures varied widely across studies, however, reports with 10 or 15 outcome measures therefore had 10 to 15 times as much influence on the results as studies containing a single outcome measure. One cannot clearly interpret the results from such reviews. Moreover, multiple effects from the same intervention are not independent, raising problems for statistical analyses. Now, the recommended practice is to average multiple effects to produce a single effect per study per analysis. In this way, each study contributes one independent data point in each analysis.†

Ruling Out Rival Explanations

Meta-analysts often distinguish between methodological and substantive features of reviewed studies. The former refer to how well studies were conducted (e.g., their experimental rigor or quality; see below), whereas the latter refer to variables that are frequently of prime theoretical, clinical, or applied importance (e.g., types of treatments, clients, and problems). What treatment is more effective? Which types of therapists obtain the best outcomes? Which problems respond most favorably to intervention? Yet, a meta-analyst cannot assume that substantive features are the only factors that influence outcomes. Similar to an individual experiment, the meta-analyst should rule out plausible rival hypotheses for the obtained results.

*Weighted effects are lower than unweighted effects when a negative correlation exists between sample size and effect size. Researchers often have difficulty in securing large samples in clinical studies, but the fewer the participants, the less reliable the findings. Therefore, weighting procedures such as Hedges and Olkin (1985) give greater weight to effects from studies with larger samples.

†Ways to deal with dependent data are beyond the scope of this chapter; see Cooper and Hedges (1994) for details.

Apart from the direct influence of substantive variables, there are four primary alternative explanations for variability in ESs: (1) sampling error; (2) study artifacts; (3) methodological features; and (4) confounded study characteristics. Therefore, beware of meta-analyses that concentrate only on substantive variables without controlling for or examining these other factors. In other words, it is an empirical question in each meta-analysis, of how much of the variability in outcomes is accounted for by sampling error, study artifacts, and various methodological and substantive features and confounds among study characteristics.

Two examples provide some perspective on this issue. In his review of 397 delinquency studies, Lipsey (1992) found that 27% of the variance in outcomes was attributed to sampling error, 15% to study artifacts, 25% to method variables, and 22% to characteristics of the treatments (11% of the variance was unaccounted for). In a review of child therapy, Weiss and Weisz (1990) found that methodological factors explained more outcome variance than treatment-related variables. As a result, before succumbing to the temptation to conclude that substantive variables are solely responsible for the obtained effects, other sources of influence must be examined.

It is possible to estimate the degree of sampling error present in effect sizes (see Hunter & Schmidt, 1990, for examples). If most of the variance in ES is due to sampling error, then there may be little point in searching for other moderators of outcome because there is not much variability among effects left to explain. Homogeneity tests are another way to determine if it is reasonable to assume that sampling error is primarily responsible for the variability of obtained effect sizes. See Hedges and Olkin (1985) for a basic explanation and Durlak, Fuhrman, and Lampman (1991), Lipsey (1992), and Shadish et al. (1993) for examples of this procedure.

Study artifacts generally refer to different sources of error in the primary studies such as unreliable outcome measures or imperfect implementation of the intervention. Hunter and Schmidt (1990) describe 11 potential study artifacts. It may not be possible to correct or control for all study artifacts, but their existence should be noted. For instance, there is a positive relationship between the reliability of an outcome measure and its corresponding ES. Because measures are not perfectly reliable, however, meta-analyses underestimate ESs due to measurement error. ESs can be corrected when the reliability of outcome measures is known using a simple formula (each ES divided by the square root of the measure's reliability), but reliability data are often missing in the primary studies.

The impact of methodology is not consistent across research domains and should be tested in each meta-analysis. Some meta-analyses have found that better designed studies yield lower ES (Ottenbacher & Petersen, 1985), while others report the opposite finding (Weiss & Weisz, 1990). Although there is no "gold standard" when it comes to methodological rigor, reviewers frequently focus on such basic design issues as random assignment to conditions, types of control groups, and what types of outcome measures are used. Other potentially important design features may be suggested by prior theory and research.

Some meta-analysts impose methodological criteria *before* studies are included in the review. For example, Shadish et al. (1993) only included marital and family studies in which subjects were randomly assigned to conditions and were clinically distressed (as opposed to including quasi-experimental designs or analogue or prevention studies).

Confounds often exist between methodological and substantive features or among the latter variables. For instance, it is not uncommon to find that some interventions are evaluated using more rigorous or sensitive experimental procedures, one type of outcome measure is used more frequently to evaluate certain treatments, or that treatments are not uniformly applied to all types of problems. A good meta-analysis will probe for such confounds before presenting final conclusions and interpretations.

Different strategies can be pursued to estimate the impact of methodological features and probe for confounds. In the former case, a system can be developed to rate studies in terms of experimental quality and these ratings can then be correlated with ESs to see if there is a significant relationship.

Alternately, studies can be divided into different categories on the basis of one or more methodological features (e.g., studies using or not using random assignment to conditions, or studies of good versus poor methodological quality) and the ESs for these categories can be compared (see Dush, Hirt, & Schroeder, 1989; Forness & Kavale, 1996, for examples of these approaches).

One can also probe for possible confounds by subdividing studies in different ways to see if other variables also account for the results. For instance, Durlak and Wells (1997) found that the ways populations were selected for intervention and the level of intervention that was conducted (e.g., identifying those undergoing important life transitions and intervening with children directly versus changing their school or home environment, respectively) were important determinants of ESs in preventive mental health programs. However, outcomes also were evaluated for other variables (such as methodological variables, intervention techniques, characteristics of the change agents, and age of the target population) to rule out these plausible alternative explanations for program outcomes.

One effective strategy to probe for confounds is through a multiple regression analysis using ES as the criterion variable. Methodological and clinical variables can be entered as possible predictors to assess their relative influence. Regression controls for the overlap (confounds) among variables and can help clarify the unique contribution of variables to outcome.

For example, Shadish et al. (1993) initially found significant outcome differences in family therapy studies favoring behaviorally oriented treatments. The results of a multiple regression analysis, however, indicated that only two methodological features (i.e., the source of the outcome measure and whether the study was published or not) were significant predictors of ES. Treatment orientation did not emerge as a significant predictor in the regression. In this case it appears that methodological factors were more important than the type of family treatment.

As another example, McGlinchey and Durlak (1997) hypothesized that comorbidity would be the most important clinical factor affecting outcomes. In a regression analysis, several design features were entered first as a block to predict ES (e.g., such variables as assignment to conditions, use of a normed outcome measure, use of an attention-placebo control versus a waiting-list or no-treatment condition). These variables explained 16% of the variance, but the next best predictor of ES was comorbidity, as hypothesized, which accounted for an additional 6% of the variance in ES.

Because the results of a regression can capitalize on chance, confidence in the findings is increased when the analyses are guided by *a priori* hypotheses about the relative ordering of variable, or if the results are replicated. If a sufficient number

of studies exist, the regression can be repeated on randomly divided halves of the total study sample. In the absence of a priori hypotheses or replication, readers should accept the results of regression analyses tentatively. Furthermore, in general, the results of a meta-analysis that fails to assess the impact of methodological features and possible study confounds should be viewed cautiously.

Is the Meaning of Outcomes Sufficiently Explained?

A mean ES is simply a number whose meaning has to be explained by the meta-analyst. It is often helpful to describe the types of outcomes on which the mean is based. Therefore, the overall mean ES that should have been initially obtained by averaging multiple ESs within studies later should be disaggregated or broken down into different outcome categories. For instance, it is commonly found that interventions for children and adolescents produce greater changes on measures of overt behaviors and self-reported symptoms than on assessments of academic performance and sociometric status (Casey & Berman, 1985; Durlak & Wells, 1998; Beelmann et al., 1991). Durlak and Wells (1997) reported that primary prevention mental health programs were equally successful in significantly reducing problems and in significantly promoting competencies in child and adolescent target populations. Disaggregated ESs are often essential in clarifying in what ways participants changed after intervention.

Is the Practical Significance of Outcomes Assessed?

It is important that the meta-analysis assess the practical or social significance of outcomes. There is no straightforward relationship between the magnitude of an effect and its practical significance; much depends on the nature of the outcome measure. A high ES (e.g., 1.0) takes on a different practical meaning depending on whether it is based on serious antisocial behavior, on-task classroom activity, sociometric status, academic performance, and so on. At the same time, a mean ES of "only" 0.20 can represent cost-effective treatment and significant practical social benefit if based on rates of recidivism or imprisonment among criminal or delinquent populations (Lipsey & Wilson, 1993; Lösel, 1995). There are now several ways to calculate the clinical significance of effect sizes (see Baucom & Hoffman, 1986; Durlak et al., 1991).

For example, Durlak et al. (1991) calculated a "normative effect size" (NES) by selecting data from different studies that used normed outcome measures. In these calculations, the data for the normative group were substituted for the control group when the ES was determined. Normative data were available for such measures as the Achenbach Child Behavior Checklist, the Connors Teachers Hyperactivity Scale, the Piers Harris Self-esteem Scale, and so on. In contrast to a customary ES, lower values are preferred for an NES to demonstrate less of a difference between the treated group and the normal comparison group. Cognitive behavioral therapy was successful in reducing the overall mean NES from pre to post (mean NES of 1.55 and 0.50, respectively), but results did vary for different types of outcome measures; much more change was noted on self-reported symptomatology relating to anxiety and depression than on measures of overt behavior. As a result,

it was concluded that treatment had produced a practical impact on adjustment, but that further behavioral improvement was still possible and desirable.

Shadish et al. (1993) analyzed their data to determine whether couples who were clinically distressed at the beginning of treatment were functioning within normal limits following intervention. Success in this respect was achieved for 41% of treated couples but in none of the control groups over time, an indication of the practical effects of marital treatment.

Readers also can gauge practical effects on their own by calculating the value of the binomial effect size display (Rosenthal & Rubin, 1982) if it is not provided in the meta-analysis: Simply convert the reported mean ES to a Pearson r by dividing by 2. Expressed in proportions, the success rate for the intervention and control groups can be found by using the following formula: $.50 \pm r/2$, respectively. For instance, Durlak and Wells (1997) found that primary prevention programs were equally successful in reducing externalizing symptomatology and enhancing children's academic achievement (identical mean ESs of 0.30). These means convert to comparative success rates of 57.5% for the intervention group versus 42.5% for controls on each outcome measure (i.e., a mean ES of $.30/2$ equals an r of .15 and then inserting this value into the equation $.50 + r/2$; the obtained proportions can be directly converted into percentages). If intervention had no effect, the success rates for both groups would equal 50%. Although a 57.5% success rate for the prevention group in this example seems low, it is 35% higher than the success rate for controls (15/42.5). Differential success rates can be calculated for any type of outcome measure. If the outcome involves life or death, which it sometimes does in medical trials, mean ESs as low as .07 can nevertheless constitute an effective intervention (Rosenthal, 1991).

Are Findings and Conclusions Appropriately Qualified?

The generality of meta-analytic findings is important. Do the results of a therapy review apply to all forms of treatment, all types of therapists, and all forms of presenting problems, or only to certain subcategories of these dimensions? Generalization of findings cannot be assessed unless the primary characteristics of reviewed studies are described in the meta-analysis.

At the very least, it is important to know some specifics about child clients (age, race and ethnicity, nature and severity of presenting problems), their families (socioeconomic status and marital status), therapists (experience and training), the treatments (major ingredients, level of implementation, duration and frequency of sessions), and outcome measures (reliability and validity).

Information about study characteristics is essential in generalizing the results of a meta-analysis. Wampold et al. (1997) reported that different types of treatments did not yield significantly different results, but they did not provide specifics on what treatments were evaluated. Crits-Christoph (1997) pointed out that only three child or adolescent treatments and no family therapies were evaluated by Wampold et al. (1997). Obviously, Wampold and co-workers' (1997) conclusion about the equivalence of outcomes among different treatments cannot be extended to child, adolescent, or family therapies.

Virtually all meta-analytic findings must be qualified in some way because of limitations in the data set:

A likely scenario with applied social science data is that there will be many studies from which one can glean some of the desired data, a moderate number from which one can retrieve most of the data, and only a few from which one can obtain virtually all of the data. (Nurius & Yeaton, 1987, pp. 701-702).*

In a methodological review of over 500 child therapy outcome research we (Durlak, Wells, Cotten & Johnson, 1995) found that researchers rarely presented details on any of the following: clients' specific social-cognitive abilities, coping styles, or levels of academic functioning, the nature and level of parental psychopathology or their child-rearing techniques, measurement of the treatment process, and data on treatment implementation. In some cases, even basic information relating to the age, gender, and problems of the study sample are missing (see Lösel, 1991). As a result, some potentially important analyses cannot be done and subsequent analyses must be appropriately qualified. For instance, we could not evaluate an important six-fold conceptual scheme prominent in primary prevention studies because of small cell sizes for many categories (Durlak & Wells, 1997). The data were also insufficient to assess how program implementation affected outcomes in reviews of both primary and secondary prevention (Durlak & Wells, 1997, 1998). The characteristics of reviewed studies must be considered when offering meta-analytic conclusions.

Are Specific Directions for Future Research Offered?

A good meta-analysis not only tries to explain prior research but also offers specific guidelines to improve future investigations. Another reason for specifying which limitations exist in the current database is that such information logically suggests how to improve future studies. There may be additional needs to compare different theories, collect certain types of outcome information, study more diverse client samples, assess how well treatment is implemented, and so on. Meta-analysts should pinpoint what potentially important information is absent and encourage future researchers to begin filling the void. For instance, Durlak and Wells (1997) suggested 14 ways to improve the theoretical, methodological, and empirical aspects of primary prevention research.

LEARNING MORE ABOUT META-ANALYSIS

The reader who wishes to learn more about meta-analysis should first read Light and Pillemer (1984) and Wolf (1986), and then some combination of Hunter and Schmidt (1990), Durlak and Lipsey (1991), Durlak (1995), Cooper and Hedges (1994), and Rosenthal (1995). These sources provide good overviews of meta-analysis and plenty of examples and procedural details. One good way to learn is by doing. Before attempting a meta-analysis on one's own, however, it would be helpful to seek consultation from those experienced in meta-analysis (e.g., authors of recently published meta-analyses). Software is available to handle the basic

*A meta-analysis always should contain details on the procedures and results of coding procedures to indicate how reliably the information of interest was captured.

statistical analyses (Johnson, 1989), but consultants can offer valuable guidance on such issues as the framing of research questions, identifying relevant studies, how to develop coding schemes that will most effectively capture the necessary data, and how much time and effort must be invested to complete the proposed meta-analysis successfully.

CONCLUDING COMMENTS

This chapter has discussed seven useful criteria for judging the quality of a meta-analysis. These criteria are not exhaustive and some exceptions may apply, but in general the more a meta-analysis satisfies these criteria, the more likely it is that the review has been carefully conducted. It is important to stress, however, that strengths in one area do not compensate for deficiencies in other areas. For instance, a biased search of the literature, inappropriate statistical analyses, or failure to qualify one's conclusions seriously compromises the overall value of the research synthesis, regardless of how well other aspects of the meta-analysis have been accomplished. Therefore, each major step in a meta-analysis is important.

REFERENCES

- Abrami, P. C., Cohen, P. A., & d'Apollonia, S. (1988). Implementation problems in meta-analysis. *Review of Educational Research, 58*, 151-179.
- Baucom, D. H., & Hoffman, J. A. (1986). The effectiveness of marital therapy: Current status and application to the clinical setting. In N. Jacobson & A. Gurman (Eds.), *Clinical handbook of marital therapy* (pp. 597-620). New York: Guilford.
- Beelmann, A., Pfingsten, U., & Lösel, F. (1991). Effects of training social competence in children: A meta-analysis of recent evaluation studies. *Journal of Clinical Child Psychology, 23*, 260-271.
- Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98*, 388-400.
- Cooper, H., & Hedges, L. V. (Eds.). (1994). *Handbook of research synthesis*. New York: Russell Sage Foundation.
- Crits-Christoph, P. (1997). Limitations of the dodo bird verdict and the role of clinical trial in psychotherapy research: Comment on Wampold et al. (1997). *Psychological Bulletin, 122*, 216-220.
- Dickersin, K. (1997). How important is publication bias?: A synthesis of available data. *AIDS Education and Prevention, 9*, 15-21.
- DuPaul, G. J., & Eckert, T. L. (1997). The effects of school-based interventions for attention deficit hyperactivity disorder: A meta-analysis. *School Psychology Review, 26*, 5-27.
- Durlak, J. A. (1995). Understanding meta-analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 319-352). Washington, DC: American Psychological Association.
- Durlak, J. A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin, 110*, 204-214.
- Durlak, J. A., & Lipsey, M. W. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology, 19*, 291-332.
- 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*, 115-152.
- Durlak, J. A., & Wells, A. M. (1998). An evaluation of indicated preventive intervention (secondary prevention) mental health programs for children and adolescents. *American Journal of Community Psychology, 26*, 775-802.

- Durlak, J. A., Wells, A. M., Cotten, J. K., & Johnson, S. (1995). Analysis of selected methodological issues in child psychotherapy research. *Journal of Clinical Child Psychology, 24*, 141–148.
- Dush, D. M., Hirt, M. L., & Schroeder, H. E. (1989). Self-statement modification in the treatment of child behavior disorders: A meta-analysis. *Psychological Bulletin, 106*, 97–106.
- Forness, S. R., & Kavale, K. A. (1996). Treating social skill deficits in children with learning disabilities: A meta-analysis of the research. *Learning Disability Quarterly, 19*, 2–13.
- Hansen, W. B. (1992). School-based substance abuse prevention: A review of the state of the art in curriculum, 1980–1990. *Health Education Research, 7*, 403–430.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Hunter, J. E., & Schmidt, F. L. (1990). Methods of meta-analysis: Correcting error and bias in research findings. Newbury Park, California: Sage.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficiency of psychological, educational, and behavioral treatment. *Confirmation from Meta-Analysis American Psychologist, 48*, 118–209.
- Johnson, B. T. (1989). *DSTAT: Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Erlbaum.
- Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. L. Light, T. A. Louis, & F. Mosteller (Eds.), *Meta-analysis for explanation* (pp. 83–127). New York: Russell Sage Foundation.
- Lösel, F. (1991). Meta-analysis and social prevention: Evaluation and a study on the family—Hypothesis in developmental psychopathology. In G. Albrecht & H. Otto (Eds.), *Social prevention and the social sciences: Theoretical controversies, research problems, and evaluation strategies* (pp. 305–332). Berlin: Walter de Gruyter.
- Lösel, F. (1995). Increasing consensus in the evaluation of offender rehabilitation: Lessons from recent research syntheses. *Psychology, Crime and Law, 2*, 19–39.
- Matt, G. E. & Navarro, A. M. (1997). What meta-analyses have and have not taught us about psychotherapy effects: A review and future directions. *Clinical Psychology Review, 17*, 1–32.
- McGlinchey, K. A. & Durlak, J. A. (1997). Therapy outcome research on child and adolescent internalizing disorders. Unpublished manuscript.
- Nearpass, G. L. (1990). Counseling and guidance effectiveness in North American high schools: A meta-analysis of the research findings. *Dissertation Abstracts International, 50*, 1948A.
- Nurius, P. S., & Yeaton, W. H. (1987). Research synthesis reviews: An illustrated critique of “hidden” judgements, choices, and compromises. *Clinical Psychology Review, 7*, 695–714.
- Ottenbacher, K. J., & Petersen, P. (1985). The efficacy of early intervention programs for children with organic impairment: A quantitative review. *Evaluation and Program Planning, 8*, 135–146.
- Prout, H. T., & DeMartino, R. A. (1986). A meta-analysis of school-based studies of psychotherapy. *Journal of School Psychology, 24*, 285–292.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (rev. ed.). Newbury Park, CA: Sage.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192.
- Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*, 166–169.
- Shadish, W. R., Montgomery, L. M., Wilson, P., & Wilson, M. R., Bright, I., & Okwumabua, T. (1993). Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology, 61*, 992–1002.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*, 752–760.
- Swanson, H. L. (1996). Meta-analysis, replication, social skills, and learning disabilities. *Journal of Special Education, 30*, 213–221.
- Tabak, E. R., Mullen, P. D., Simons-Morton, D. G., Green, L. W., Mains, D. A., Eilat-Greenberg, S., Frankowski, R. F., & Glenday, M. C. (1991). Definition and yield of inclusion criteria for a meta-analysis of patient education studies in clinical preventive services. *Evaluation and the Health Professions, 14*, 388–411.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta analysis of outcome studies comparing bona fide psychotherapies: Empirically, “All must have prizes.” *Psychological Bulletin, 122*, 203–215.
- Weiss, B., & Weisz, J. R. (1990). The impact of methodological factors on child psychotherapy outcome research: A meta-analysis for researchers. *Journal of Abnormal Child Psychology, 18*, 639–670.

- Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, 55, 542–549.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, 117, 450–468.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.

19

Reviewing and Editing Manuscripts for Scientific Journals

DENNIS DROTAR

Reviewers and editors of manuscripts for articles that summarize scientific research that is published in scholarly journals perform several critical functions for the field, most especially that of gatekeepers of scientific knowledge. In this capacity, they are charged with determining which manuscripts are of sufficient quality that they should be published and disseminated to other scientists versus those that should not be published because they do not contribute new scientific knowledge and/or have significant methodological problems.

The second important function of editors and reviewers is to improve the quality of scholarly writing and science of those manuscripts that are eventually published. Reviewers and editors provide critical feedback to authors to help improve the clarity of their writing and the quality of their research. Finally, editors and reviewers help to determine the priorities and policies of a particular journal, such as the type of articles are emphasized and the boundaries of a journal's interests.

Although reviewers and editors are important, if not critical, to the quality of the science that is published in a particular field, they obtain relatively little direct training to perform their tasks, unless they have practiced conducting reviews under supervision as a part of their training. Consequently, many reviewers develop their skills through "on-the-job" training only after they have been selected as a reviewer or member of an editorial board. For this reason, I believe that the tasks of editing and reviewing not only require more attention in graduate

DENNIS DROTAR • Department of Pediatrics, Rainbow Babies and Children's Hospital, Cleveland, Ohio 44106-6038.

and postdoctoral training, but greater description and explication in professional writing.

My experience as a reviewer for more than 20 years on the editorial boards of the *Journal of Pediatric Psychology* (JPP) and the *Journal of Clinical Child Psychology* (JCCP), as associate editor for JPP from 1990 to 1997, and ad hoc reviewer for many journals focused on research with children has generated "lessons learned" about the structure and process of reviewing and editing scientific articles. These are summarized in this chapter. My hope is that this information will be helpful to authors in pediatric and clinical child psychology who submit manuscripts as well as to reviewers and editors who review them. Readers who are interested in this topic might wish to consult the following articles: Council fo Biology Editors Peer Review Retreat Consensus Group (1995), Goldbeck-Wood (1998), Maher (1978), Roberts, Lyman, Breiner, & Royal (1982), Schwab (1985), and Squires (1989, 1990). Moreover, the chapter on preparing research articles for publication in this volume also contains information concerning the manuscript review process and should be read as a companion piece to this chapter.

OVERVIEW OF THE STRUCTURE OF AN EDITORIAL BOARD OF A SCIENTIFIC JOURNAL

The work of a scientific journal is accomplished by a group of scientists, led by the editor, that has the responsibilities of reviewing manuscripts that are submitted for publication in the journal's content area. The editorial board of a journal includes associate editors, who along with the editors manage the review of individual manuscripts, and reviewers, who are assigned by the editors to review individual manuscripts. Although the specific responsibilities of editorial board members may vary somewhat from journal to journal, the central task of editorial board members involves reviewing individual manuscripts. In order to manage the volume of manuscripts that are submitted most journals have a relatively large number of editorial board members who are necessary to process manuscripts without unduly taxing any one reviewer. The availability of a large number of editorial board members also facilitates the editor's ability to match the specific content areas of submitted manuscripts with reviewers' areas of expertise. For example, in 1997, JPP had three associate editors and 53 members of the editorial board, while the JCCP had two associate editors and 60 consulting editors. In addition to the editorial board members, most editors also rely heavily on the work of ad hoc reviewers who are selected to review individual manuscripts in special areas of their expertise. Ad hoc reviewers give editors increased capacity to match reviewer expertise with the varied content of manuscripts.

EDITORS' ROLES AND RESPONSIBILITIES

The editor, who is charged with the responsibility of ensuring that journal operations run smoothly, is responsible to the governing body of the professional organization that sponsors the scientific journal and selects the editor. Editors will select several associate editors to help them manage the work of a journal. These

associate editors function as managing editors for individual manuscripts, provide consultation to the editor as needed, and help shape editorial policy. The editor also has the responsibility of processing the manuscripts as they are accepted for publication and the organization of each issue, including the order of manuscripts in the journal.

Managing the Review Process

The editors and associate editors of a journal who are the managing editors for individual manuscripts are responsible monitoring the overall progress of journal reviews, including timeliness and quality. They select reviewers for a particular manuscript and monitor the progress of the reviews, including reminders to tardy reviewers.

The managing editor has the important task of deciding whether a manuscript has the potential to make a significant scientific contribution. Editors make three basic decisions about the disposition of the manuscript: acceptance, revise and resubmit, or reject (see Chapter 16, this volume). In making these decisions, editors need to distinguish between potentially correctable problems in the writing and/or methods contained in a manuscript versus those that are not. While almost every manuscript can be improved by revision, not every manuscript warrants one. Consequently, the editor must determine the following: (1) if the ultimate contribution of the manuscript to the field is significant enough to warrant the considerable effort and time involved in revisions, and (2) what specific revisions are necessary to improve the manuscript. In my experience, editors can be especially helpful by identifying "diamonds-in-the-rough," that is, well-executed studies with potentially interesting data that are packaged in manuscripts that lack focus and clarity, and by working with authors to improve the quality of such manuscripts.

After they have reviewed the manuscript and rendered a decision, a managing editor writes a disposition letter to the author that summarizes the substance of editorial feedback and advises him or her about the next steps, for example, whether to revise the manuscript or submit to another journal. This disposition letter serves as the primary source of guidance to the author. Because some reviewers may disagree about the relative merits of a manuscript and weigh various aspects of a work very differently, editors face a considerable challenge in integrating reviewers' opinions and delivering a clear summary statement to the author.

Setting Editorial Policy and Communicating with the Editorial Board

Competent editors are often proactive in shaping a journal's policy by setting standards for manuscript acceptance and selecting the topic areas or themes for special journal issues that may be managed by other scholars with specialized expertise who are selected for this task. For example, the special or topical issue has proved to be a popular format for readers and authors of JPP and JCCP. Recent examples of their special issues include the April 1997 issue of JPP, edited by Anne Kazak, that was devoted to family systems issues in pediatric psychology and the March 1998 issue of JCCP, edited by Sandra Russ, that focused on developmentally based integrated psychotherapy with children.

Another critical set of editors' responsibilities involves communication with members of the editorial board, including associate editors and reviewers. Competent editors are in active dialogue with reviewers and consulting editors to manage dilemmas raised by individual manuscripts as they occur and to obtain advice to improve journal policies and the process of review. For example, editors will inform reviewers if they are late with a review and also may provide feedback to individual reviewers concerning the quality of their reviews.

Relationship of the Editor to the Organization That Sponsors a Journal

The editor of a professional journal is appointed by an organization that sponsors the journal [e.g., the executive council of a section or division of the American Psychological Association (APA) to carry out the responsibilities of the editorship]. Editors are appointed to their position, generally for a 6-year term, after a thorough search process in which his or her professional qualifications, including stature as a scientist, editorial experience, letters of recommendation, and an applicant's statement of future goals for the journal are carefully reviewed.

Once the editor is selected, his or her performance is reviewed by the governing body of the organization. Although this happens rarely, it is possible for the editor to be removed from their position in the case of problematic performance. Many editors are in close communication with the executive committee or board of the sponsoring organization and meet regularly with them to discuss issues, problems, and policies in the journal's operation. Editors generally submit an annual report, which includes a detailed summary of the journal's operations, to the sponsoring organization's governing body.

Deciding to Become an Editor

The editor's time consuming job includes screening manuscripts for relevance to the journal's mission, deciding on reviewers, reviewing manuscripts, reading and synthesizing reviewers' comments, writing letters that summarize feedback and editorial decisions to authors, handling inquiries and correspondence from authors communicating with reviewers, and so forth. The motivation for editors to conduct these tasks is intrinsic. Although journal editors and associate editors are paid a stipend by the publishing company, most of this pays for supplies and secretarial help and does not begin to cover the time that is spent in conducting editorial business.

Given the extraordinary nature of editorial work, a critical question for prospective editors is: Why would I want to take on the responsibilities of reviewing and editing? The answer to this question is a highly personal one. In his retrospective, "Confessions of an Editor," Routh (1995) described his personal motivation for his distinguished editorial work (as editor of JPP, JCCP, and now *Journal of Abnormal Child Psychology*) as stemming from his family tradition of editorial and scholarly experience and his strong commitment to encourage students and younger colleagues in effective self-expression. Those of us who have clearly benefited from his feedback, mentorship, and counsel can attest to the strength of the Routh family tradition. My own motivation for editorial work has stemmed in

part from a strong interest in helping to enhance the quality of published work, to help shape the direction of research, and to provide consultation to authors.

EDITORIAL BOARD MEMBERS' ROLES AND RESPONSIBILITIES

Manuscript Review

Members of the editorial board who conduct reviews of manuscripts for scientific journals are the heart and soul of the process of peer review. They are charged with developing and implementing the standards by which manuscripts are reviewed. The major responsibilities of the editorial board are to contribute timely and thorough reviews of manuscripts.

In carrying out these tasks, reviewers serve several masters: the editor, the author of the manuscript, and the journal's readership. Responsibilities to the editor include adhering to instructions for reviews and the time lines for reviews (depending on the journal, the required turnaround time is 3–6 weeks); providing a balanced, thorough review to guide the editor's decision; and giving detailed, constructive feedback to the author concerning the manuscript (Council of Biology Editors Peer Review Retreat Consensus Group, 1995). Reviewers' responsibilities to both the author and editors are to conduct an impartial, unbiased review and to preserve the confidentiality of the manuscript. The journal's readership is served by the reviewers' maintaining high scientific standards for published work.

It is important to note that reviewers' multifaceted roles are not easy to reconcile. For example, a reviewer's role as gatekeeper of science necessitates a comprehensive critique that carefully documents the strengths and weaknesses of a manuscript, including why it should or should not be published. This role is not easy to reconcile with the responsibility to give constructive feedback in order to improve the quality of the author's scientific contribution. An ideal resolution of such role strain is for the reviewer to provide a detailed, objective assessment of the scientific quality of a manuscript's contribution along with a constructive review. This is no small accomplishment and requires specialized scientific knowledge, empathy, tact, and precision in writing. Consequently, reviewers who are able to consistently achieve such a standard are highly valued by editors and authors.

Over and beyond their contributions in reviewing manuscripts, editorial board members contribute their ideas concerning journal policy, scholarly activities that they want to promote (e.g., case studies, ideas for special issues), and ways to improve the operating efficiency of the journal. Consequently, the quality of a journal's operations is much improved by the active involvement of the editorial board.

Selection of Editorial Board Members

Editorial board members are selected for a time-limited term on the basis of their scholarly contributions in the content areas that are represented in the journal's statement and purpose. They are selected by editors who have knowledge of their scholarly work or have received recommendations from other scientists, especially members of the editorial board, or other editors. Because it is impossible

for editors to have direct acquaintance with the scholarly work and reviewing skills of a large number of scientists, editors rely heavily on others' recommendations to select reviewers.

One characteristic but untested assumption in selecting reviewers is that successful experience and achievement in publishing manuscripts translates directly into the ability to provide a balanced, thorough, and timely critique of manuscripts. Based on my personal experience with editorial boards as editor and reviewer, this assumption is upheld more often than not. On the other hand, as those of us who have received limited, inaccurate, and otherwise less than helpful reviews can attest, eminence in a particular area of research provides no iron clad guarantee that a reviewer will deliver timely, high quality reviews (Nelson, 1996).

Feedback to Editorial Board Members

While editors strive to develop a way to provide feedback to members of the editorial board, in practice it is in fact quite difficult to provide direct feedback to reviewers concerning the quality of their reviews for several reasons: Reviewers volunteer their time and editors rely on them heavily. It is also difficult for editors to determine how to give feedback, especially negative feedback, to their reviewers. Detailed feedback to reviewers also would be inordinately time consuming. Consequently, editors' feedback to reviewers tends to be general and positive. Editors can provide an important positive reinforcement to reviewers for a job well done by giving detailed, constructive feedback to authors. Reviewers who have spent the time and energy necessary to provide high-quality reviews appreciate such feedback concerning their work.

The main way that editorial board members receive feedback is by comparing their reviews with those of other reviewers and the editor's letter of decision to the author. As a reviewer, I have learned a great deal by comparing my reviews with that of other reviewers, (e.g., areas of a manuscript that I may have glossed over or missed alternative ways of presenting information to authors, etc.). For these reasons, I would encourage reviewers to carefully review all editorial correspondence they receive.

Some journals (e.g., the *Journal of Behavioral and Developmental Pediatrics*) also provide summaries to reviewers concerning their individual response times and those of other reviewers blinded. This same journal provides certificates of appreciation for reviewers who, in the editors' judgment, have consistently contributed timely, thorough, and frequent reviews. Moreover, JPP recently has given similar acknowledgments to reviewers.

Making the Decision to Join an Editorial Board

When asked to be on the editorial board of a journal, one's initial reaction may be a reflexive: "sure, I'd be happy to." It is flattering to be asked because the invitation is a sign that one's work is recognized by one's colleagues. Moreover, being named to an editorial board is helpful for one's career advancement, as it is a benchmark of peer recognition. Reviewing also provides a way to keep in touch with new scientific work and maintain professional contact with colleagues who are on the editorial board.

Despite the above advantages of editorial board membership, I would recommend that prospective reviewers carefully consider the demands that are required before they accept this responsibility, lest they make commitments they cannot keep. Depending on the number of board members and the number of submissions, members of editorial boards may review as many as 10 manuscripts per year for a single journal. Considering the time spent in reading and thinking about the manuscript and composing and editing the review, a single review may take anywhere from 1 to 4 hours per manuscript. Moreover, reviewers are “on call” in the sense that they need to respond promptly. When one considers that many reviewers are members of multiple editorial boards and also give time to other journals as ad hoc reviewers, the time involved in editorial work can be formidable and involve a considerable amount of unpaid compensation. Such unselfish donation of time and energy is becoming more and more difficult to make, in light of multifaceted professional pressures faced by reviewers, especially those pediatric and clinical child psychologists who see patients, teach, and conduct research.

Given the time, energies, and skills that are necessary to provide high-quality reviews of manuscripts, potential reviewers should ask themselves two questions before agreeing to serve on an editorial boards: (1) do I really want to be a reviewer, considering the time, energy, and skill requirements? and (2) do I have the time to provide competent reviews within the time frame that is required? If the answer is no to either one of these questions, a prospective reviewer may want to decline the invitation.

CHARACTERISTICS OF EFFECTIVE REVIEWS

The hallmarks of an effective review are thoroughness, clarity and specificity, constructiveness, tact, and timeliness (Goldbeck-Wood, 1998; Hyman, 1995).

Thoroughness

What is meant by a thorough review? This involves nothing less than a careful reading of the entire manuscript, together with carefully drawn, point-by-point evaluative comments for each section of the manuscript. Excellent reviewers go that extra mile by reviewing each section of the manuscript, supplying information to support points that are made in their critiques, and elaborating their points clearly and carefully. Some reviewers will also supply references that may be useful to authors. Others provide detailed, line-by-line feedback, which also can be extremely helpful to authors.

Clarity and Specificity

Clarity of feedback in a review is also prized by authors and editors for obvious reasons. Reviewers are not immune to problems of unclear writing. An unclear and/or nonspecific review does not give sufficient guidance to authors and leaves it up to them to try to figure out the points that are made. For example, a statement such as “the data analyses were problematic” does not give the author precise

information about what the problem was, let alone information about how to improve it.

To enhance the clarity and specificity of their reviews, reviewers should make liberal use of specific examples from the manuscript to document specific points or sections that were seen as problematic. For example, reviewers who are concerned about the data analysis section of a manuscript should be explicit about the nature of their concerns. Explicit, detailed comments such as “the rationale for the order of entry of the variables for the regression analysis reported on page 7 was not clear and should be clarified” are more helpful to authors than general statements about the inadequacy of their approach to data analysis. Similarly, sentences or sections that reviewers found difficult to understand should be clearly indicated either on the manuscript itself or by indicating the specific pages and lines that were not clear.

While providing a high level of specific detail, including detailed examples, in reviews is much more time consuming than general feedback, it is well worth it in the service provided to authors and editors. A lack of specificity in reviews not only is frustrating to authors but makes it difficult for them to know how to improve their manuscript. Most authors have received some reviews that caused us to shake our heads and wonder: “what is he or she talking about?” When feedback from editorial reviews cannot be understood, even after discussions with colleagues, I would encourage authors to contact the managing editor to obtain clarification rather than risk misinterpreting content of the review.

Constructiveness

Constructiveness is another critical but elusive dimension of an effective manuscript review. Some reviewers are better at finding flaws in a manuscript or study design than they are at recognizing its essential strengths and potential to contribute to the scientific literature. Moreover, in their zeal to identify problems, some reviewers can easily neglect the strengths of a manuscript. Nevertheless, a clear and comprehensive description of a manuscript’s strengths can be important to the editorial decision. Moreover, by giving credit where it is due, reviewers craft a more credible review.

In my experience, some reviewers provide excellent critiques of the strengths and weaknesses of a manuscript but have more difficulty in articulating constructive suggestions for improving the manuscript. To help concretize the abstract dimension of “constructiveness,” it may be helpful for reviewers to remember how they felt as authors when they received reviews that provided a litany of criticisms but little or no feedback concerning how problems might be addressed to make the manuscript more acceptable. For example, in the case of the ambiguous regression analysis referred to earlier, the author might be instructed to provide a detailed rationale for the order of entry of variables based on a theoretical model.

Constructive feedback takes considerable skill because it is usually easier to identify what is wrong with a manuscript than it is to suggest strategies to address the problems that can be rectified. At the same time, authors should recognize that reviewers are not obliged to help them address each and every problem with their manuscripts. After all, it is a review, not a research-related consultation. Moreover, there is a threshold beyond which it is not realistic for reviewers or editors to

provide comprehensive feedback, such as when articles have to be totally rewritten or when the scientific contribution of a study is judged to be marginal (Routh, 1995).

Tact

Reviewers and editors are called on to deliver difficult information about a manuscript respectfully and in a way that can be heard by authors. A simple “golden rule” of reviewing would read: write your review in a way that you would like to receive for your own manuscripts or the way that you would write a trusted colleague who has consulted you on a research problem (Nelson, 1996). It may be helpful to re-read one’s reviews from the standpoint passing “the tact test.” A tasteful review can and should have incisive criticism but at the same time not contain language that is demeaning or implies “you should have known better.”

While reviewers should not be concerned about pulling their punches in noting problems in manuscripts, they should appreciate the potential impact of their choice of words and the tone of their reviews. Consider the differences in the following comments on the same problem. The first is more critical: “The author should have recognized the need to adjust the alpha level to correct for type 1 error in the analysis. This is a careless mistake.” The second comment is more straightforward and tactful: “In order to guard against type 1 error, alpha levels need to be adjusted.”

Timeliness

Timeliness is a critical feature of a high-quality review. Professional psychologists have multiple pressing responsibilities that can easily take precedence over the task of reviewing a manuscript. How widespread a problem is the delay in reviews? Nelson (1996) described her experience with reviewers’ turnaround time in her tenure as the editor of *Cognitive Development*. In contrast to her goal, which was to have reviewers respond within 3 weeks, only 15% of editorial board members sent their replies within the time limit. About 16% said they could not review the manuscript and returned it within several weeks. Almost half (53%) of the editorial board reviewed the manuscript between 1 and 3 months. Some of the reviewers responded only after being sent a reminder letter. Another 12% responded after the 3-month period after multiple letters, phone calls, and “special pleading.” A small group (4%) never responded (Nelson, 1996).

On the other hand, my direct experience with editorial board review in managing manuscripts as associate editor for the JPP was more encouraging, although there were always some reviewers who were late and others who did not respond to reminders. For example, the average lag for reviews for JPP for 1997–1998 was only about 6 weeks.

Nelson (1996) attributed the problem of late manuscript reviews to several factors such as overuse of some reviewers, using more (e.g., 3–6) reviewers for a single manuscript, using multiple rounds of review and revisions, or involving different reviewers in a revised manuscript who were not initially consulted. Nelson (1996) suggested that editors may aggravate the problem of delays in reviews either by delaying to write a decision letter, not resolving conflicting editorial

opinions, or giving unwarranted encouragement for a revision to manuscripts that ultimately are not accepted, even after multiple revisions.

TYPES OF EDITORIAL FEEDBACK

In general, reviewers' feedback deals with two basic areas of the manuscript: (1) scientific design and contribution (e.g., the quality of the research design and the contribution to new knowledge); and (2) the quality of writing (e.g., how the research was presented). In addition, the content of feedback from reviewers can be classified as: (1) major questions/concerns (e.g., defined as methodological problems judged to significantly limit the scientific contribution of the report); and (2) minor concerns or questions that involve such issues as incomplete or ambiguous descriptions of methods. While some studies are rejected because of specific fatal flaws (e.g., lack of a control group in an intervention study), in my experience the editorial decision is more likely to reflect the cumulative impact of both major and minor concerns about the study and its presentation. For this reason, it is helpful to editors and authors for reviewers to clarify how they weighed the various problems that they identified in the manuscript and to indicate whether they consider these problems as major or minor flaws. Such feedback helps authors to determine whether and how they should revise a manuscript when they are given a "revise-and-resubmit" editorial verdict. Major editorial concerns about method or questions about the overall level of contribution of the manuscript may not be correctable or worth correcting, depending on the quantity and nature of the problems. On the other hand, problems that involve clarity of writing or need for more detail in the context of an otherwise solid design are correctable, albeit time consuming.

Communication to Authors about the Editorial Decision

The editor's disposition letter to the author should clearly describe the editorial decision and advice to the author concerning the next steps concerning the manuscript. The tone of the editor's letter may be particularly important in the case of a revise-and-resubmit editorial verdict, which may leave authors wondering whether they should resubmit the article or submit to another journal. Editors cannot guarantee eventual acceptance of an article that has been revised after receiving a revise-and-resubmit editorial verdict and will inform authors of this fact. At the same time, editors will be reasonably encouraging in their letter if they feel that there is scientific merit in the manuscript and/or if the manuscript addresses an important problem that has not received much empirical scrutiny.

ETHICAL ISSUES IN REVIEWING AND EDITING MANUSCRIPTS

Editors and reviewers of professional journals have with important ethical responsibilities that include ensuring integrity and accuracy of the editorial process, confidentiality, and limiting potential conflicts of interest (Chusid, Casper, & Camitta, 1984; Maher, 1978; Morgan, 1985).

Ensuring Quality and Accuracy of Editorial Review

One primary ethical responsibility of reviewers is to ensure that their reviews are the highest quality and that they make accurate and fair statements concerning the content of the manuscripts that are reviewed. Moreover, reviewers and editors have the responsibility to check accuracy and inconsistency in the scientific data that is presented (Cullitan, 1987). At the same time, there are obvious limits on reviewers' and editors' abilities to function as watchdogs for scientific accuracy, especially in the case of scientific fraud (Keith-Spiegel & Koocher, 1985).

Confidentiality of Reviews

Manuscripts under review are to be regarded as privileged and confidential communication (Cullitan, 1987). As such, they must be seen only by reviewers and editors and not circulated to anyone else. Potential ethical problems in reviewing manuscripts involve such actions as reviewers discussing a manuscript with someone else or circulating a manuscript under review to another colleague who the reviewer feels may be "interested" in it. While sharing information and manuscripts with colleagues is a well-accepted practice in the scientific community, manuscripts under review are not in this category. Reviewers may want to consult colleagues in their setting who have specialized expertise on a topic related to the manuscript. If they choose to do so, they should inform the editor and credit their colleagues.

Another potential ethical problem is raised by reviewers who give a manuscript to another person, often a junior colleague, to give them experience in reviewing. While this practice can provide useful opportunities for training in manuscript review, it should be cleared by the editor and involve direct supervision by the senior reviewer (see subsequent section on training in manuscript review).

Managing Conflicts of Interest in the Review Process

Another set of ethical issues that can arise in the review process concerns conflicts of interest between reviewers and authors. For example, a reviewer may feel as if they cannot review the manuscript of a former student or close professional associate as impartially as they otherwise would. Moreover, reviewing work from one's arch rival or enemy could engender negative bias. Knowledgeable editors can prevent these problems if they know about them, but they cannot always identify them (Routh, 1995).

Reviewers have the ethical responsibility to recognize a potential for conflicts of interest and to excuse themselves from any review if they feel that such bias could interfere with an impartial review. On the other hand, because the fields of pediatric and clinical child psychology contain a relatively small group of researchers, one inevitably finds oneself reviewing the work of colleagues. The question that only reviewers can answer and editors need to monitor is whether their relationship (positive or negative) with the author is such that it precludes a reasonably impartial review. While blind editorial review may protect against such conflicts to some extent, experienced reviewers are often able to recognize the work

of colleagues who conduct research in areas related to their own work (Fisher & Friedman, 1994).

Another set of ethical issues are posed by manuscripts in which the editor or associate editor is the author. In such cases, it may be necessary to appoint someone else to the position of managing editor who assigns reviewers and who renders the editorial judgment for such manuscripts. However, even if these submissions are handled fairly, as evidenced by the fact that editors do have their work rejected by their own journals, there can still be the appearance of undue influence by the editor (Routh, 1995). For this reason, some editors may choose to minimize submissions to the journal that they are editing.

IMPROVING THE MANUSCRIPT REVIEW PROCESS

Most journal editors are concerned with developing ways of improving the review process for research and reviewers. Epstein (1995) presented several useful suggestions for improving the review process, which were made by editors and reviewers. Some of these and other suggestions are reviewed below.

Exercise Independence of Editorial Judgment

Epstein (1995) recommended that editors should be instructed that their primary task is to exercise their own editorial judgment on the basis of all the evidence that they are given. For this reason, editors should be willing to override the recommendations of reviewers in the event of sufficiently compelling cases that are made by authors.

Identify and Control Editorial Biases

Reviewers and editors should be aware of their own preferences for theories and research paradigms and recognize that such biases can influence their decisions concerning the manuscripts they review. Epstein (1995) argued that editors should avoid selecting reviewers who are known to share the same theoretical views, especially when the manuscript under consideration threatens a prevailing model. In the event of serious disagreements between authors and reviewers in an area beyond the editor's expertise, the editor should feel free to consult other experts for an additional opinion.

Providing a Meaningful Appeals Procedure

Epstein (1995) also advocated for journals to have a standard appeals procedure for authors who believe that their rejected articles have been improperly evaluated. This would include a mechanism by which someone other than the original editor would review the previous review to determine whether a request for re-review should be honored.

Enhancing Feedback to Reviewers

One of Epstein's (1995) most interesting suggestions is that editors should provide authors with standard forms for evaluating the reviews that they receive, for example, to rate the degree to which they regard a review as conscientious, constructive, and/or objective. In my view, such data would give reviewers much needed feedback concerning their reviews and give a voice to the all-important consumers of manuscript reviews: the authors.

Ways to Enhance Timeliness of Reviews

Most journal editors struggle with how best to reduce the turnaround time of the review process. While there are no easy answers to improving the lag time for reviews, reviewers, editors, and authors share some responsibility to reduce the waiting time for authors, as described in the next section.

Reviewers' Responsibilities

In order to enhance the turnaround time of reviews, editorial board members should place a reasonable priority on manuscript reviews, which clearly compete with their other responsibilities. For example, heavily used reviewers with a particularly strenuous workload might want to set aside time for editorial work in their work week. In the event that a manuscript is not sent, they can utilize the time for other work. Nelson (1996) suggests that reviewers reserve having a "special pile" in their correspondence for incoming reviews.

If a manuscript review will be delayed, it is necessary to inform the editor when you will be able to get to it (Nelson, 1996). As a managing editor, I have very much appreciated reviewers contacting me to let me know if they are going to be late with their review and when they anticipate that they will complete their assignment. To be most efficient, reviewers should not wait for a pile of manuscripts to accumulate. Returning a manuscript immediately if it cannot be reviewed within the time limit that is requested is the best policy, because it facilitates the editorial assignment of new reviewers.

Editors' Responsibilities

Editors can facilitate the timeliness of manuscript reviews if they carefully monitor the progress of manuscripts and contact tardy reviewers to remind them to send in their review as soon as possible. In my experience, most reviewers respond to the editor's personal reminders about review deadlines. Editors also can reduce reviewers' overall time commitment by refraining from recommending a revise-and-resubmit editorial verdict for manuscripts that are marginal in their potential scientific contribution and/or would take an extraordinary amount of revision and subsequent review (Routh, 1995). While such decisions are difficult, they can enhance the efficiency of the review process and reduce the author's frustration concerning manuscripts that are ultimately rejected. Moreover, editors should not consult reviewers on each and every manuscript revision but only on those revisions that are extensive and/or are especially difficult to review (Nelson, 1996).

Authors' Responsibilities

Finally, authors also should take some responsibility to influence the timeliness of the editorial process by contacting managing editors if they have not heard anything about their manuscript within the stated time line for editorial review. Such reminders can help to jog dilatory editors and reviewers into appropriate action.

To Blind or Not to Blind Reviews

One of the more controversial areas concerns the issue of whether the reviewer should be blinded as to the identity of the authors. Different editors have different opinions on this issue and policies differ from journal to journal. For example, the JPP has a blinded review policy while the JCCP does not. Available empirical data suggest that blinding does not appear to make a significant difference in the nature of the editorial decision (Fisher, Friedman, & Strauss, 1994), partly because it is sometimes possible to guess the identities of authors even when they are anonymous (Fisher & Friedman, 1994). Nevertheless, my own personal preference as a reviewer is not to be informed about the identity of the person whose work I am reviewing.

Another difficult issue is whether the reviewers should be identified to authors. In the interest of protecting reviewers from potentially irate authors, most journals do not inform authors who the reviewers are. On the other hand, Epstein (1995) advocates that reviewers should be identified to authors to reduce the operation of bias. The counterargument here is that reviewers' anonymity may help them to honestly report what they believe. It is interesting to note that some journals, for example, the JDBP, gives reviewers a choice concerning whether they want to remain anonymous. If they do not wish to preserve their anonymity, the names of all reviewers are noted on the editor's correspondence to the reviewers. As a reviewer for Journal of Developmental and Behavioral Pediatrics, I must say I am even more alert to the need for a thorough review because my review can be so readily compared with those of my colleagues.

TRAINING IN MANUSCRIPT REVIEW

Formal training in reviewing manuscripts review can be helpful for several reasons: First, it familiarizes students and/or junior colleagues with the process of manuscript review. Such experience also can help students develop a reviewer's perspective about their own writing, especially to their appreciation of the value of clear writing. Finally, the experience of evaluating a manuscript, which includes putting one's criticisms and suggestions in writing in a constructive manner, can be helpful in learning to give effective critique to others.

I have used several different methods to facilitate graduate students' training in reviewing manuscripts. One vehicle that I have found to be helpful is a workshop format in which students engage in mutual critique of one another's manuscripts and other writing projects (see Chapter 16, this volume). Specific training in manuscript review can be implemented in this format if students are also asked to

provide written critiques of one another's manuscripts. These written critiques can then be evaluated from the standpoint of clarity, thoroughness, tact, and so on.

Another approach I have used is to give students supervised experience in preparing critiques of articles that have been published and/or accepted. Students are given a manuscript, such as one of my own, that has already been through the review process and has been published. Students who prepare critiques of these manuscripts have the opportunity to compare their critiques against those of experienced reviewers and the editor. In situations where the manuscript received a revise-and-resubmit verdict, student reviewers can follow the review process and evaluate the authors' response to the critiques and the disposition. Such involvement in following a manuscript review from start to finish gives students a good feel for the process of reviewing and an opportunity to see how manuscripts are improved by the review process.

Another vehicle for providing training in reviews is to give students or junior colleagues the practice of generating reviews under "gamelike" conditions by giving them opportunities to review an article that was sent to the senior or supervising reviewer. I am not particularly fond of using this as a teaching method because I do not like to incur additional time demands for myself (e.g., editing the students' reviews while also performing the reviewer's task), which could delay the review. Moreover, some students have difficulty responding to the strenuous demands of a timely review. On the other hand, as an editor I have seen this process work well in hands of the right combination of supervisor and student reviewer.

I would recommend several safeguards for reviewers who work with student reviewers. The student's participation in the review should be cleared with the editor who had sent the manuscript to a particular reviewer and is expecting a review from that person. Moreover, the reviewer of the manuscript should not simply turn the manuscript over to the student, but should closely supervise the review. So long as the review reflects the joint product of the supervisor and junior reviewer's work, the author and editor can reap the benefit of both perspectives. Finally, students who conduct reviews under supervision should get proper credit for their reviews in the form of acknowledgment in the journal's list of ad hoc reviewers.

BECOMING A REVIEWER FOR PEDIATRIC AND CLINICAL CHILD PSYCHOLOGY JOURNALS

If they have the time and interest to be a reviewer, pediatric and clinical child psychologist researchers should be assertive and volunteer themselves for this task. In my experience, editors are quite interested in recruiting new, energetic, competent researchers to participate in the editorial process. In fact, it is useful that editors have some experience with reviewers before they are appointed to an editorial board, as well as for reviewers to have firsthand experience in what is involved in a review. Consequently, I would encourage interested researchers to volunteer their services as reviewers by sending a curriculum vital and statement of interest to the editor(s) of the journal(s) of your choice. You have little to lose and only manuscripts to gain!

REFERENCES

- Chusid, M. J., Casper, J. T., & Camitta, B. M. (1984) Editors have ethical responsibilities too. *New England Journal of Medicine*, 331, 990–993.
- Council of Biology Editors Peer Review Consensus Group. (1995) Peer review guidelines: A working draft. *Council of Biology Editors Views*, 18, 79–81.
- Cullitan, B. J. (1987). Integrity of research papers questioned. *Science*, 235, 422–423.
- Epstein, S. (1995). What can be done to improve the journal review process. *American Psychologist*, 50, 883–885.
- Fisher, M., & Friedman, S. B. (1994). The blind reading the blind. *Journal of Developmental and Behavioral Pediatrics*, 15, 386–387.
- Fisher, M., Friedman, S. B., & Strauss, B. (1994). The effects of blinding. *Journal of the American Medical Association*, 272, 142–146.
- Goldbeck-Wood, S. (1998). What makes a good reviewer of manuscripts. *British Medical Journal*, 316, 86.
- Hyman, R. (1995) How to critique a published article. *Psychological Bulletin*, 118, 178–182.
- Keith-Spiegel, P., & Koocher, G. P. (1985). Ethics in psychology. Professional standards and cases. Hillsdale, NJ: Lawrence Erlbaum.
- Maher, B. A. (1978). A reader's writer's and reviewers guide to assessing research reports in clinical psychology. *Journal of Consulting and Clinical Psychology*, 46, 835–838.
- Morgan, P. P. (1985). An author's code of ethics for medical editors. *Canadian Medical Association Journal*, 132, 731.
- Nelson, K. (1996, Spring). Observations of an editor on professional responsibilities. *SRCD Newsletter*, 1–3.
- Roberts, M. C., Lyman, R. D., Breiner, J., & Royal, G. P. (1982). *Publishing child-oriented articles in psychology. A compendium of publication outlets*. Washington, DC: University Press of America.
- Routh, D. K. (1995). Confessions of an editor, including mistakes I have made. *Journal of Clinical Child Psychology*, 24, 236–241.
- Squires, B. P. (1989). Descriptive studies: What editors want from authors and peer reviewers. *Canadian Medical Association Journal*, 141, 879–880.
- Squires, B. P. (1990). Statistics in biomedical manuscripts: What editors want from authors and peer reviewers. *Canadian Medical Association Journal*, 142, 213–214.

VI

Developing Strategies to Integrate Research and Practice

One of the critical needs for the fields of pediatric and clinical child psychology research concerns more effective integration of research and practice. Traditionally, research and practice have been relatively isolated in these fields. For example, much of the research on treatment of childrens' behavioral problems and mental disorders has been conducted on populations that are not typically seen by practicing clinicians. Moreover, the overwhelming majority of psychological research that has been conducted on children with chronic health conditions has focused on description of psychological status rather than interventions to modify these problems. Consequently, the generalizability and the practical relevance of available research to practitioners and policy makers who are concerned about the delivery of services for children and families with psychological disorders has been limited. However, the problems that are faced by researchers who are interested in conducting clinically relevant research on populations that are seen by practitioners are formidable. The strategies that need to be used to conduct research on psychological services that are delivered in communities are different from those that are taught in graduate school. Consequently, researchers need to be taught specialized skills to implement clinically relevant research. Equally significant challenges are faced by researchers who are interested in designing research programs that are closely tied to relevant clinical issues and practitioners who want to use research to guide the development of their programs. To provide information and models for readers to master the formidable task of integrating research and practice, each of the contributors to Section VI focuses on a topic that relates to the promotion of better integration between research and practice. In Chapter 20, Linscheid describes relevant issues in designing and describing case studies that continue to be an important but neglected vehicle to and practice in the fields of pediatric and clinical child psychology. The advantages of single-subject designs are described along with the various types of case studies that are possible for researchers to report. Linscheid takes the reader step by step through the process of describing and presenting a sample case report and describes what is necessary to document the impact of treatment on a single case, including type of effect, stability of the therapeutic effect, number and heterogeneity of subjects, and identification and elimination of logical alternatives for the treatment effect. Linscheid also gives

several useful recommendations to enhance students' training in conducting case reports and case series.

To address the need to develop clinically relevant research that is conducted on real-life problems and populations, the Chapters 21 and 22 are contributed by experienced researchers who have developed research programs to evaluate treatments and interventions that are conducted in community settings and in clinical populations. In Chapter 21, Henggeler and Randall describe the results of their research programs to evaluate the efficacy of randomized trials of multisystemic therapy. Drawing on experiences that have been used in conducting randomized trials of multisystemic therapy, the authors focus on issues of internal validity that are often neglected by researchers including randomization, treatment specification, and ensuring treatment integrity. Neglected aspects of external validity including sample generalization, research attrition, treatment retention, and inter-agency collaboration are also discussed. In considering how randomized clinical trials can be conducted in community settings, Henggeler and Randall identify and describe the critical tasks for researchers who are conducting randomized trials of treatment in community settings (gaining collaboration from community agencies in research mental health services, ensuring treatment specification, and developing strategies that are necessary to promote and enhance treatment integrity across different settings). Finally, Henggeler and Randall also consider practical strategies of assessing participant characteristics, recruitment and retention and issues in minimizing rates of treatment drop out that are necessary to conduct high quality research on the efficacy of treatments for children and adolescents in implementing research concerning children's mental health services.

In Chapter 22, based on their extensive experience in conducting research concerning mental health efficacy of treatment, Douglas, Kelley, Nixon, and Bickman consider issues in evaluating mental health services for children and adolescents in community settings. These authors clarify distinction between treatment efficacy and effectiveness, underscore the importance of program theory and the necessity for theory-based measurement, and highlight the need to monitor implementation of the research program. Recommendations such as using pilot studies in developing critical information about the population of interest are made to facilitate researchers' ability to conduct research in community settings. Moreover, the authors consider the issue of referred populations and issues of developing research collaborations with multiple parties and the interpretation of findings. Implications for methods to enhance the clinical relevance of mental services research are also described.

One of the important ways to promote the clinical relevance of research in pediatric and clinical child psychology is to develop integrated programs of research and clinical practice in practice and research settings. The essence of such integration involves research that is informed by the needs of practitioners and practice that is informed and shaped by research findings. The integration of research and practice is a deceptively simple concept that is very difficult to operationalize and implement. Chapters 23 and 24, each of which were developed and composed by experienced researchers, describe tactics and strategies of integrating research and practice in the fields of pediatric and child clinical psychology. In Chapter 23, Blount, Bunke, and Zaff describe integration of research and practice in a research program focused on research in pediatric pain. Using a

behavior analytic paradigm, Blount and colleagues illustrate how clinical practice provides a model of how clinical research helps to determine factors that influence target behaviors. A model to integrate explicative and treatment research is presented based on the authors' research that focuses on the antecedents and consequence of acute procedural pain in children. Blount and his colleagues' detailed presentation of a program of research provides useful insights concerning how research and practice can be integrated, for example, by direct observation and a thorough knowledge of the patient population.

In Chapter 24, Kazak and Meadows describe the comprehensive integration of psychosocial research and practice within a specialized program for psychosocial services for children with cancer and their families in a pediatric hospital setting. The authors describe the theoretical underpinnings of their research program in a framework of social ecology. Within this program, research questions have been guided by clinical issues and concerns that have emerged in the course of management of children's cancer, for example, psychosocial sequelae of childhood cancer survival, pain and distress during medical procedures, and neuropsychological outcome in bone marrow transplantation. Kazak and Meadows describe the clinical implications of their research and the critical importance of interdisciplinary collaboration to develop an integrated program of research and clinical care in a hospital setting. Moreover, their suggestions for developing and maintaining an integrated program of research and practice, for example, strengthening collaborative links with hospital administration, framing clinical services within a clear conceptual model, using research to provide evidence of effectiveness of services, and developing multiple sources of funding provide models for others who wish to follow their lead.

20

Case Studies and Case Series

THOMAS R. LINSCHEID

Despite concerns about scientific rigor and many potential interpretive problems, case reports and case series have played a major role in shaping the field of clinical psychology. The purpose of this chapter is to define the role of the case study in relation to other types of clinical research, to delineate types of case studies, and to review criteria to be used in evaluating the validity of conclusions drawn from case studies.

While seeming to be the least sophisticated and least labor intensive of the methods of clinical research, the case study can be difficult to conduct in applied settings and requires attention to special strategies and methods in order to increase the validity of conclusions that may be drawn. Because other research strategies such as single-subject or group research designs may require design restrictions not in concert with "real-world" situations, the case study is sometimes the only way to provide clinically relevant information; therefore, if the case study is conducted with procedural considerations to be described in this chapter, it can yield meaningful clinical information. Case studies can be meaningful in their own right but they can also serve to shape research questions that may be answered by use of the more sophisticated designs.

Perhaps the most dramatic example of the impact that case studies and case series have had on the field of clinical psychology is the development of effective treatments for phobias using the reciprocal inhibition model. Watson and Rayner's (1920) description of fear conditioning in the famous "Little Albert" case report (methodological and ethical concerns notwithstanding) suggested that fear-anxiety could be learned, laying the groundwork for the development of counter-conditioning interventions. The important and influential case series report of Wolpe (1958) suggested the effectiveness of this approach, and subsequently more sophisticated group intervention research designs have verified the effectiveness of these techniques.

THOMAS R. LINSCHEID • Department of Psychology, Children's Hospital, Columbus, Ohio 43205.

Case reports and case series also have proven to be important in the origin and establishment of specialty areas of research or application. Drotar, La Greca, Lemanek, and Kazak (1995) describe the importance of case reports in the development and legitimization of the field of pediatric psychology. They document the predominance of case reports in the early volumes of the *Journal of Pediatric Psychology* and discuss how these reports shaped the focus of the field by defining research and practice agendas. Interestingly, as the field grew, case reports in the journal declined in number, making way for more sophisticated group research designs that were primarily descriptive and explicative in nature (Roberts, 1992).

While it seems scientifically desirable for a field of study to move to increasingly more sophisticated research, Drotar and colleagues lament the decline in case reports in the *Journal of Pediatric Psychology* which they feel has led to an increasing separation between research and practice. As a side note, Drotar (personal communication) reports that there has been little response to the call for submission of case study reports. This emphasizes the concern that case studies, more often conducted in the applied world, will not be viewed as legitimate research by the academic world.

Hayes (1983) also expresses concerns that emphasis on group research designs has "helped drive a wedge between research and clinical practice" (p.181). Because clinical practice is individual case based, group research designs often separate the clinician from input into a legitimate knowledge base in the field and produce research that may not be applicable in individual treatment situations. Hayes argues that methodologically sound case analysis, rather than being a "weak sister" in clinical research, should be at the core of the clinical research endeavor. Jones (1993) agrees that traditional psychotherapy research has had little impact on clinical practice and argues, in a special section of the *Journal of Consulting and Clinical Psychology*, that single-case research deserves more respect for its ability to provide clinically meaningful information about psychological treatment.

Recently, the Association for Advancement of Behavior Therapy, recognizing the problem of the researcher-clinician split, began publishing a journal titled *Cognitive and Behavioral Practice*. It is designed to bridge the gap between researcher and clinician; the organization also has offered to "arrange marriages" between practicing clinicians and researchers to facilitate the writing of case report manuscripts.

CONTRASTING CASE STUDIES WITH GROUP AND SINGLE SUBJECT DESIGNS

In the process of discovery, of which scientific research plays an important role, there will always be the issue of the accuracy of results versus the importance of results. In psychology, the acceptance of research findings has traditionally been influenced by attempting to determine the probability that the observed results occurred by chance and by reducing the probability that the results are attributable to another cause. In traditional group research, statistical methods are used to assess the probability that results were obtained by chance variation and experimental controls (randomization, counterbalancing, etc.) are used to decrease the chance that results are related to factors other than those being manipulated.

However, to attain an accurate estimate of the probability of occurrence by chance and to document an effect using group designs, certain sacrifices must be made. In order to attain experimental control of relevant variables, research questions may be narrowed and consequently lose clinical relevance. That is, the variables of interest may not be relevant to those observed in clinical practice. Second, populations under study may not be adequately described or selected because of difficulties in randomization and documentation of all the possible components of a population that may be related to outcome. Also, the need for statistical analysis emphasizes group over individual outcomes and statistical significance rather than clinical significance (Wells, 1987).

Time series-based single-subject research designs have been developed to introduce experimental control so as to increase the probability that observed changes in behavior are truly the result of the manipulated intervention (cf. Barlow & Herson, 1985). Treatment withdrawal, reversal, multiple baseline, and alternating treatment designs can be used to decrease the probability of an observed behavior change being the result of factors other than those manipulated by the experimenter. While exact probabilities are not determined as in a group statistical analysis, single-subject designs can establish that alternative explanations for behavior change have such low probabilities that they can be discarded. While group statistical analysis are not integral components of single-case experimental designs, statistical procedures have been developed to assign probabilities to observed behavior changes in relation to chance (cf. Busk & Marascuilo, 1992).

Single-subject designs have the advantage of addressing clinically relevant questions with a high degree of confidence that results are attributable to intervention manipulations. Disadvantages include the need for intense person power to collect behavioral observation data over time and the potential lack of generalizability.

The case study or case series report, though lacking sophistication experimentally, offers information that may be of significant importance despite the absence of procedures to verify statistically or probabilistically the accuracy of the result. Case reports are primarily differentiated from group designs and single-subject research designs on the degree of control of extraneous variables and the extent of repeated observations. They offer the ability to provide information in numerous circumstances in which group or single-subject designs are not possible. For example, single-subject research designs often require observation of baseline rates of behavior. The baseline provides a comparison rate of behavior against which to assess the effects of an intervention. However, if the behavior is dangerous to self or others, it may not be practically or ethically possible to allow the behavior to occur freely in order to ascertain the rate prior to initiation of treatment. In addition, withdrawal of an effective treatment as a means of demonstration of clinical control may not be possible if the behavior is dangerous or there is reason to believe that clinical control may not be reestablished.

In practice, certain clinical problems or phenomenon occur so rarely that group designs are not possible. The rarity of a problem may preclude procedures of random assignment and introduces time-based variables if groups are formed over extended time periods. Certain statistical procedures require minimum sample sizes, and therefore do not lend themselves to small-group comparisons or analysis of individual treatment results.

TYPES OF CASE STUDIES

Yin (1993) divides case studies into three types: exploratory, descriptive, and explanatory. The exploratory study is used to define new questions and provide hypotheses for further study. This type of case study was used by Freud and his followers in the development and expansion of his theory of psychosexual stages of development. Descriptions of clinical cases and phenomenon were analyzed for underlying themes, and from this structures of analysis derived. The literature also contains reports of parent's reflections on certain issues such as the death of a child (Coolidge, 1977) and parents' experiences in advocating for effective but controversial treatments (Van Duser & Phelan, 1993). These reports most often contain little objective information but do provide insight into the subjective experiences of those receiving clinical services. These subjective reports may serve to define research area in which higher-level research designs can be utilized to document the commonality of these experiences and patterns of adjustment. Descriptive case studies are used to illustrate a diagnostic or theoretical issue. They are not concerned with clinical or therapeutic behavior change, but rather they provide actual or anecdotal data to describe a phenomenon. A common type of descriptive report in clinical psychology may describe the intellectual or emotional functioning of an individual or individuals with a rare disorder. For example, Mansheim (1979) reported emotional and behavioral data obtained from a boy with a rare genetic disorder. Likewise, Sassman, Zartler, and Mulick (1981) provided descriptions of cognitive functioning in two sisters with a low-incidence biochemical disorder. The literature in clinical child and pediatric psychology contains reports that are both exploratory and descriptive. Most typically, such reports present a theoretical or conceptual model of a clinical problem or approach to treatment and include the application of the model to an example case. For example, Titter & Cook (1981) describe the development of a systems framework for working with families and schools and then present the case of a 10-year-old student with learning disabilities to illustrate how the model is applied. In this combined type of study the emphasis is on the model or conceptualization, and the case is presented more as an example than as an explanatory report; hence, the criteria for evaluating the case report itself may be somewhat less stringent than in pure explanatory case studies.

Explanatory case studies, concerned with cause-effect issues most often related to treatment outcome, are perhaps the most recognizable in clinical psychology research and often hold the most potential for application by individual practicing clinicians. In addition, they may provide impetus for the development of treatment approaches, which may later be evaluated using larger scale, more sophisticated research designs.

It would be safe to say that most case studies are published by practitioners who, while aware of the requirements for scientific validity, find themselves in a situation where immediate therapeutic effects outweigh considerations of experimental control, reliability of measurement, and so forth. As a somewhat absurd example, consider the practitioner who agrees to see a child with a recent history of fire setting. The practitioner devises a treatment strategy and as a scientist-practitioner is interested in whether the proposed treatment can be established as the operative factor in the hoped-for behavior change (i.e., elimination of fire setting). To verify scientifically that the treatment is responsible for therapeutic

improvement, the therapist could withhold treatment until a large number of referrals for fire setting are made and then form groups by randomly assigning patients to the proposed treatment or to another control condition (no treatment, alternative treatment). As an alternative to the group design approach, the therapist could decide on a single-subject research design in which the child's rate of fire setting would be documented before and after treatment introduction using a number of possible single-subject research designs. Obviously, both possibilities, which would be necessary to verify scientifically that the intervention was responsible for the behavior change, are impractical because both involve allowing the fire setting behavior to continue until the requirements for scientific rigor are met. As an alternative, the clinician could choose to begin treatment with the referred fire setter and present the results as a case study. This explanatory case study would then be evaluated based on a set of common sense criteria to be detailed in the remainder of this chapter.

EVALUATING THE CASE REPORT

While ethical and logistic reasons may prevent the degree of scientific rigor needed to establish cause–effect relationships, explanatory case reports are more or less believable based on a number of factors. The probability that the reported intervention was responsible for the noted therapeutic change depends not on controlling threats to internal validity as is done in the controlled research study but rather on the probability that other factors or coincident events could have affected the behavior in question. Case reports therefore should include discussions of factors that could pose threats to internal validity and any evidence that the author has to suggest that possible alternative explanations are improbable.

Characteristics of the case study itself also affect believability. Kazdin (1992) describes the following dimensions for evaluating case studies: type of data, assessment occasions, past and future projections, type of effect, and number and heterogeneity of subjects. In addition to Kazdin's considerations, case studies should be evaluated on (1) the effort to identify other possible and plausible explanations and provide evidence that these were not operative, and (2) the stability of the treatment effect. These dimensions will be examined in relation to an example case report to illustrate how they impact on the probability that the therapeutic intervention was responsible for the reported outcome.

EXAMPLE CASE REPORT

Cunningham and Linscheid (1976) treated repetitive self-induced vomiting or rumination in the 9-month-old infant. The behavior began when the infant was 5 months old and a number of medications had been unsuccessfully used to treat the behavior. In addition, the child had undergone extensive medical testing with no significant findings. Nearly a month prior to onset of treatment, the infant was hospitalized and frequent hydration with intravenous fluids was required to sustain life. The existing psychiatric literature at that time suggested the use of treatments involving increased stimulation and affection based on the hypothesis

that the behavior developed as a self-stimulatory mechanism to compensate for decreased stimulation and affection from the mother or caregiver. With the infant's life in jeopardy, the infant's physician approached the authors with a request to use contingent electric shock to treat the ruminating behavior.

A previously published case report (Lang & Melamed, 1969) described the successful treatment of an infant ruminator with contingent electric shock, suggesting that such a treatment may be successful with the referred infant. Ruminations were counted for one baseline day and treatment was initiated the next day. Treatment consisted of administering a brief electric shock immediately for each instance of rumination. Treatment effects were documented by recording the patient's weight and number of ruminations per day (Figs. 1 and 2). The rate of ruminating decreased dramatically compared to the baseline rate measured the day before treatment began. In addition, treatment was conducted in a variety of settings around the hospital where data were collected on the infant's ruminating behavior. Slight increases were noted on days 7, 10, and 13, which corresponded to treatment in these different locations. With the decrease in ruminations, weight gain occurred rapidly.

DIMENSIONS FOR EVALUATING THE CASE REPORT

Type of Data and Assessment Occasions

For explanatory case studies, evidence is presented to document that therapeutic change has occurred. This evidence or data can take many forms including narrative self-reports by a therapist or client, paper-and-pencil self-report measures, standard psychological assessment instruments (e.g., IQ tests), direct behavioral observation measures, or objective measures of outcome (products). Data of the various types are used to establish that the behavior or problem existed prior to intervention and indeed there was change in the desired direction following intervention. The plausibility that change did occur could be influenced by three factors related to the data. The first factor is the objective versus subjective nature of the data presented. Narrative self-reports are subjective and therefore are inherently subject to bias. A depressed child's statement of feeling better following treatment may be an accurate report or may be influenced by a desire to please the therapist. Second, the number of measures used can affect the plausibility that a change has occurred. If it can be documented that the depressed child who reports that he feels better also resumes favorite activities and sleeps and eats better, the probability that the depression has improved is greater. In practice, there are many clinical problems that may require the use of subjective or self-report data. Ramsden, Friedman, and Williamson (1983) report the treatment of infantile migraine headaches in a 6-year-old female using a contingency management procedure. Because actual pain and pain reports may not correlate highly, their report would have been stronger if they had documented actual behavior or activities suggestive of headache reduction (e.g., participation in sports or social activities) in addition to their documentation of reduced headache complaints. As the authors imply, using contingency management procedures for pain reports may have served to reduce the reports without reducing the true number of headaches. In

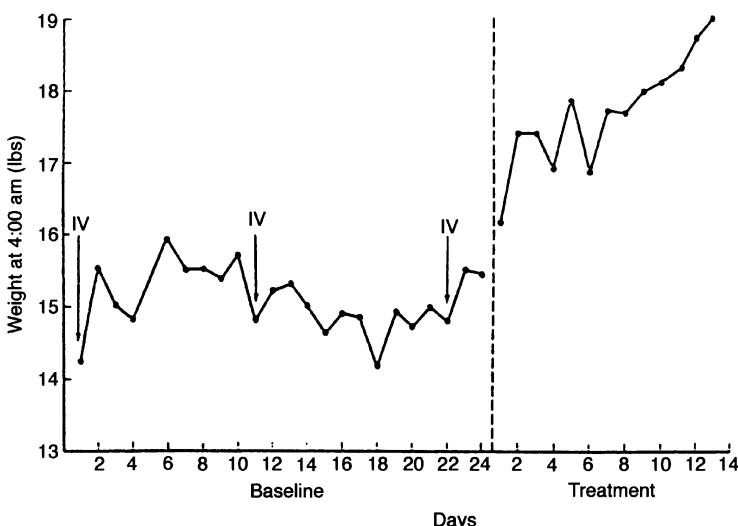


Figure 1. Patient's weight at 4:00 AM during baseline and treatment. IV indicates days on which intravenous feedings were required. Reprinted from Cunningham, C. E., & Linscheid, T. R. (1976). Elimination of Chronic infant rumination by electric shock. *Behavior Therapy*, 1, 231-234. New York: Academic Press with permission.

effect, they are able to conclude only that their treatment reduced headache complaints, not actual headaches.

Third, the greater the number of assessment occasions both pre- and posttreatment, the more likely it is that real change can be assumed. Continuous or frequent measurement provides an index of variability before treatment and reveals trends in the behavior being measured. For example, in the above example of the de-

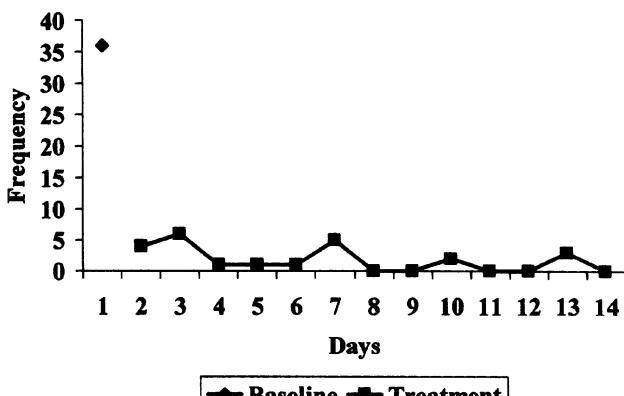


Figure 2. Frequency of ruminations per day.

pressed child, if the amount of sleep per night had been recorded for 2 weeks prior to treatment and found to vary from 1 to 8 hours per night with an average of 4 hours per night but the average sleep during the 2 weeks after treatment was found to be 8 hours, there would be greater confidence that sleep had improved. Given the variability in sleep duration observed during baseline, recording sleep duration for only one night prior to treatment may have led to a dramatic over- or underestimation of pretreatment sleep duration. This then may have led to a misinterpretation as to the effectiveness of treatment. In addition to the knowledge of sleep duration variability provided by the 2 weeks of measurement prior to treatment initiation, an increasing, decreasing, or stable pattern of sleep duration can be determined. The conclusion that treatment for depression resulted in improved sleep would be seriously questioned if the child had slept on average only 2 hours per night during the first week of the 2 week pretreatment baseline and 6 hours per night during the second week of baseline. Knowing that there was a 4-hour per night improvement in sleep duration the week prior to treatment would make it difficult to ascribe the improvement to 8 hours of sleep to the treatment. Case reports that use only one measurement occasion prior to treatment should include other measures of pertinent factors that would support the existence, extent, and possible trend of a the problem in question prior to treatment.

In the example case study, objective data in the form of daily weights (see Fig. 1) and actual counts of rumination (see Fig. 2) were used to document behavior change. Because the child had been in the hospital for over 3 weeks prior to treatment and because of his poor medical condition, it was decided to actually count the number of ruminations for only 1 day prior to treatment. A 1-day baseline measure of ruminations could be considered to be inadequate because natural daily variability in rate could lead to over- or underestimates of the rate of rumination when only 1 day was sampled. An additional measure, the patient's weight, which had remained low but stable for 24 days prior to treatment, served as an indirect measure to document that weight gain preventing rumination was occurring prior to treatment. Anecdotal reports of rumination episodes in daily nursing notes also suggested a daily rate of rumination consistent with the 1-day baseline.

The problem of absence of extended data on the occurrence of a behavior prior to intervention is common in case reports. Other sources of information, from anecdotal to objective, are required to verify that the behavior or condition existed and to serve as a baseline against which to compare treatment effects. In the example case, a by-product of rumination, the patient's weight (i.e., absence of weight gain resulting from lost food) was used. It was reasoned that if rumination decreased or ceased, weight would increase. In this way, the patient's weight served as evidence of the problem prior to treatment and as an additional means to measure therapeutic outcome.

It may be said that the more objective the data, the greater the number of measures and occasions of measurement, the stronger the case report. Narrative data, while subjective and subject to bias, can add strength to a case report if they support the objective measures. The concept of social validity (Kazdin, 1977) has led some authors to report not only objective measures of change but to include statements from parents or caregivers as supporting evidence that change had occurred and, perhaps more importantly, that the change was recognized as therapeutic by individuals other than the therapist (cf. Foxx, Bittle, & Faw, 1989).

Past Course and Future Projections

The believability of an explanatory case report is directly tied to the probability of behavior change independent of intervention and the duration of the behavior or condition prior to treatment. Continuous or frequent measures as described above are important in addressing these issues, especially in establishing the stability or trends in behavior rates. However, knowledge of the natural course of a disorder or behavioral condition also can be used to make predictions about the future and therefore to compare what would have occurred if intervention had not been undertaken. In clinical work with children, the rapid rate of development both physically and cognitively must be considered when making these predictions. The acquisition of grammatical rules and exceptions to those rules is an excellent example. Without intervention, children learn to use "went" instead of "goed" as an exception to the rule of adding "ed" to form the past tense of a verb. An intervention designed to teach children the correct rule exceptions would have to be evaluated against the naturally occurring acquisition of these rules and exceptions.

Bed-wetting or nighttime enuresis is another clinically relevant example. The designation of bed-wetting as a behavior or psychological problem depends on the child's age. Clearly, bed-wetting in an 18-month-old is expected and normal, while bed-wetting in a 15-year-old is not. While virtually 100% of children pass urine during the night before age 2; this figure is reduced to only about 15–20% percent in 5-year-old children (Walker, 1995). Barnett (1983) reports the successful treatment of enuresis in a 4-year-old girl using a behavioral procedure. Because there is a rapid and steady decrease in bed-wetting in the absence of intervention between ages 2 and 5, there is a relatively high probability of spontaneous remission of bed-wetting in the absence of treatment in a 4-year-old. The Barnett report must be evaluated against this baseline. In contrast, Noll and Seagull (1982) report successful treatment of enuresis in an 8-year-old. In this case, enuresis had persisted well past the time that spontaneous remission is likely to occur, and therefore intervention is more likely to be related to the elimination of bed-wetting.

The duration of the clinical problem affects the believability of a case report. Obviously, significant change in a problem that has been occurring for months or even years coincident with the onset of treatment reduces the chance that another factor occurred at the onset of treatment, which could provide a more plausible explanation. Additionally, a problem of long duration is less likely to change spontaneously. Conversely, clinical problems with short duration or that are tied to specific situations have a higher likelihood of changing (for better or worse) without intervention. With the increased likelihood of change without intervention comes the decreased ability to ascribe change to an intervention. Many children become situationally depressed by the loss of friends caused by a family move. A successful adjustment often occurs without intervention as the child makes new friends in school or through activities in his or her new neighborhood. Improvement in depressive symptoms following treatment in this situation may have occurred more because of the opportunities to make new friends than because of the psychological intervention. In the Barnett (1983) report mentioned above, the child treated had been dry at night for 2 years prior to the recurrence of bed-wetting associated with the birth of a sibling. In this case, the bed-wetting may have been

transitory and may have spontaneously resolved as the child and family adjusted to the new family member.

In the case example, the infant had been ruminating for 4 months prior to initiation of treatment and medical and other psychological interventions (increased contingent and noncontingent attention) had been attempted. The behavior had been occurring at home and for nearly a month's duration in the hospital and the absence of any weight gain over those months suggested that the behavior was stable. It seemed logical that, given the duration of the behavior and the lack of response to less intense treatments, rumination would likely continue if more aggressive intervention was not undertaken. From a behavioral standpoint, there was no reason to believe that the reinforcement for rumination, hypothesized to be self-stimulation, would change, also suggesting that the behavior would continue. In addition, what little research was available suggested that rumination was a chronic problem resulting in death in a significant number of cases (Kanner, 1957).

Type of Effect

Case reports are stronger when they describe "sledgehammer" effects or those that are so immediate and dramatic that it would be difficult to logically ascribe them to chance variations or some other factor occurring coincident with onset of treatment. Documenting the speed and degree of change of course is dependent on the nature of the data used to measure therapeutic change. Therefore, measures must be capable of revealing not only the degree of change, but observations must be made frequently enough to reveal when change occurred.

The slower the behavior change following the introduction of an intervention, the more probable it is that other factors may be responsible wholly or in part for the observed change. Consider the number of changes that may occur in a child's life over a several-month period. The end of a school year, the onset of allergy season, or a family move are factors that may have significant effects on a child's adjustment and may occur coincidentally with psychological intervention. Clearly, the longer it takes to observe a treatment effect, the greater the probability that other factors might be the actual reason for noted changes. The duration of the behavior or condition prior to treatment onset, as mentioned earlier, is significant in this consideration. The longer the duration of the behavior prior to treatment, the greater the likelihood that factors such as a family move or summer vacation may have occurred in the past without having a significant impact on the rate of the behavior. Teichman and Eliahu (1986) report using structural family therapy and behavioral techniques to treat two separate tics in an 11-year-old girl. The tics were extinguished in 10 to 14 weeks by author's report. Lacking knowledge of the natural course of tics, it is difficult to have extreme confidence that the described treatment was indeed responsible for the diminished rate of the tics. If tics were conceptualized as a nervous habit, factors that may have contributed to nervousness in the child may have changed unrelated to intervention (e.g., the end of the school year). In contrast, Dash (1981) reports the rapid and dramatic reduction of a needle phobia in a 5-year-old patient. Most phobias are not expected to resolve spontaneously without intervention and the speed of the treatment effect argues for ascribing the reduction of the phobia to the intervention.

Magnitude as well as speed of change is important in evaluating the results of case studies. Small changes may simply be natural variations in rate of behavior or intensity of symptoms. Aggressive children do not commit the same number of aggressive acts each day but rather show variations based on many factors including mood, opportunity, and so on. Knowledge of the actual or expected variability of a behavior or state allows for the evaluation of the magnitude of observed change against the variability of the behavior or state. The smaller the difference between normal variability and the observed change following intervention, the less believable the case study.

Unfortunately, the lack of homogeneity within psychological diagnosis makes it difficult to know the "natural course" of the problem or disorder. While medical doctors may be able to predict the course of chicken pox, for example, psychological problems rarely follow such a predictable course. However, there are logical assumptions about behavior disorders that can be made and these need to be considered when evaluating explanatory case studies.

Stability of the Therapeutic Effect

In addition to the speed of effect and the magnitude of effect, the stability and duration of the therapeutic effect may determine the believability of a case report. A study that reports follow-up data on the behavior to show that the treatment effect is enduring is a stronger report than one that reports only an initial response to treatment even if that response is dramatic and rapid. Linscheid and colleagues (1996) reported the successful and rapid treatment of choking phobia in four children using an inpatient behavior-based intervention package. In all four cases, children were able to eat solid food in normal quantities within 3 days or less. The report further documented that the children were still eating normally 3 months later. The inclusion of the 3-month follow-up data provided evidence that the treatment had most likely resulted in the elimination of the phobia. Without the follow-up data, it would have been unclear whether the children had eaten only to gain release from the hospital or had truly experienced a reduction in their fears of choking. Had they eaten only to gain release from the hospital, the effectiveness of the treatment package would be in question.

While inclusion of follow-up data to show that the therapeutic effect is lasting can serve to strengthen a case report, there are cases in which a therapeutic effect may be lost without negating the importance of a case study. As an example, consider the case in which parent training has been used successfully to treat aggressive behavior in a child. Several months later the aggressive behavior may have returned to pretreatment levels. If it can be shown that the return of aggressive behavior corresponded to parents terminating the intervention, the case study may actually be stronger. In this situation, the parent's termination resulted in a unplanned but instructive single-subject reversal design in which a therapeutic effect was observed during treatment implementation with a return to baseline levels coincident with the termination of treatment. Actually, this scenario results in stronger evidence that the intervention was responsible for behavior change than if no follow-up data had been reported.

Conversely, the believability of a case report can be negatively impacted when follow-up data are presented indicating that the behavior problem or condition has

returned even though treatment implementation has continued. When treatment continues to be implemented and the therapeutic effect is lost, the probability that the treatment was responsible for the initial change decreases.

In the previously described example case study, by the end of the first day of treatment there had been a reduction in rate of rumination from the baseline rate of 36 ruminations per day to 4. By day 3 of treatment, only one brief incident of rumination occurred and the infant had gained weight. For a condition that had a 4-month duration with evidence of steady rates of rumination per day and no weight gain, the reduction to one rumination by day 3 and the onset of weight gain could be considered dramatic and rapid and therefore not likely to have occurred by chance. In group designs the probability of observed changes can be statistically calculated; however, in case reports it is necessary to use logic and consensus to determine whether the noted change in rate or speed of change is beyond what might have occurred without intervention, assuming that other plausible explanations have been ruled out.

After 10 days of treatment, the child was discharged, and by parents' report the child did not ruminate again. The parent's report of the absence of rumination was supported by the fact that the child was severely overweight (above 90th percentile) at a 3-month follow-up visit and needed to be placed on a diet.

Number and Heterogeneity of Subjects

While the term "case report" implies a single subject or case, a number of cases are often presented in a case series. It is the absence of specific experimental manipulations to control extraneous variables that differentiates case reports, even if more than one subject is presented, from single-subject research designs or group research. Generally, the greater the number of cases demonstrating a therapeutic effect and the more similar each case is to the other in terms of response to treatment, the more believable the case report. The greater the number of cases, the less is the likelihood that an uncontrolled variable may have occurred at the onset of treatment in each case.

Heterogeneity of cases can refer to demographic issues (e.g., age, sex, and socioeconomic status), the nature of the problem, as well as to the similarity in treatment effects. The probability that the treatment was responsible for the therapeutic change is increased when it is shown that it has produced that change in children who may vary by age or other demographic factors. This refers to the robustness of the treatment effect and the probability that it is effective in more than one situation. Heterogeneity in terms of treatment success increases the believability of the report. If the observed changes in behavior were due to coincidental extraneous factors occurring with onset of treatment, the likelihood of similar effects in terms of magnitude, speed, or duration would be low.

Identification and Elimination of Logical Alternatives

Perhaps the most important factor in evaluating a case study is the extent to which the author has identified other possible explanations for the treatment effect and provided some type of evidence to show that the alternative explanations have a very low probability or have been eliminated. The identification of plausible alter-

native explanations indicates that the author is aware of other factors that may have affected the outcome of the report and has considered their impact when determining the relationship between treatment and outcome. In a simplistic example, the report of a successful psychological intervention to reduce school-related anxiety in a child with school avoidance should describe when during the school year the intervention took place. If the intervention was initiated on the last day of school, the odds are better that anxiety reduction was due to the removal of the anxiety-producing situation (school) than to the psychological intervention.

Often psychological and pharmacological interventions are undertaken simultaneously. It is not uncommon for these two interventions to be managed by different individuals (e.g., psychologist and psychiatrist) who may not be working together. Changes in medication or dose levels that occur during a psychological or behavioral intervention may be entirely responsible for observed changes, may interact with the psychological intervention, or may have no effect at all on outcome. An indication in the report that the child was not on medication or that medication was not changed prior to or during the intervention should be a standard requirement when reporting interventions for conditions that are commonly treated with medication.

Failure to accurately report information on factors that could logically affect the condition or behavior in question can have serious implications. Lack of consideration of other logical explanations or incorrectly assessing their impact can lead to erroneous conclusions about the effect of the reported intervention but also can mask important information about the effectiveness of the competing explanations. Linscheid and Landau (1993) describe a dramatic example of faulty conclusions that can be drawn when lack of consideration or inaccurate reporting of alternative explanations occurs in a report.

Berkman and Meyer (1988) published a case report suggesting that a nonaversive behavioral intervention was effective in reducing self-injurious behavior and increasing prosocial behavior in an individual living in an institution. They presented evidence that behavioral treatment programs using aversive conditioning procedures had not been successful in treating the behaviors in question, and that the behaviors had been long-standing and had been resistant to treatment with various medications. They reported that medication (Thorazine) was held constant at 1200 mg/day and that there were no major changes in the individual's living situation during the time that the nonaversive behavior intervention was implemented. Given the long-standing nature of the problem behaviors, the reported failure of pharmacological and behavioral interventions, and the stable living environment, it was logical to assume that the observed therapeutic changes that occurred during the nonaversive intervention were attributable to the nonaversive treatment program. This report was published at a time when there was significant controversy over the use of aversive procedures to treat individuals who were not able to provide consent to treatment. The Berkman and Meyer report played a significant role in the debate over the use of aversive procedures because it suggested that nonaversive procedures could be successful in treating behavior that had been shown to be resistant to treatment by both pharmacological and aversive behavioral treatment procedures.

Several years after publication of the Berkman and Meyer case report it was introduced as evidence in a legal procedure in which the prescribed use of an

aversive procedure was being contested. The lawyer representing the side proposing an aversive procedure asked for and obtained the raw data on which the Berkman and Meyer report was based. Included with the data were the medication records for the time period during which the nonaversive intervention was being evaluated. The medication records revealed that rather than the individual being on a daily dose of 1200 mg of Thorazine, Thorazine had been introduced at a dose of 300 mg/day at the same time that the nonaversive intervention was initiated. The dose then had been increased by 300 mg each per week until it reached 1200 mg. It was increased once more to 1400 mg. In addition, an effort to systematically reduce the dose had resulted in dramatically increased rates of problem behaviors despite the nonaversive intervention remaining in effect.

Because of the inaccurately reported medication dose, it had been logical to conclude that the nonaversive program was effective. In reality, decreases in problem behaviors with increasing doses of medication followed by increases in those problem behaviors corresponding to reductions in medication constituted a single-subject reversal design and provided strong evidence that the behavior change was attributable to the medication rather than the nonaversive intervention.

While it was never determined how the medication doses came to be incorrectly reported in the Berkman and Meyer article, it is clear that the accuracy of descriptions of plausible alternative explanations can significantly affect interpretation of results. In this case, based on the original report, medication as a determinant of behavioral improvement appeared to be ruled out; however, when accurate information about the medication was revealed, medication became the most logical explanation for the behavioral improvement.

CLINICAL REALITIES AND CLINICAL EFFECTS

While it may seem logical that case studies would be fairly simple to conduct, in reality these studies in contrast to group or single-subject studies are made difficult by a number of factors. Most prominent among these factors is the fact that they are conducted on actual clinical cases in applied settings. Unlike group or single subject research studies in which a specific treatment is planned for a specified clinical problem, case studies are usually conducted on cases that present first and then intervention is planned. The inability to plan treatment, decide on exclusion criteria, and recruit subjects, in contrast to group studies, means that the subjects of case studies are not as likely to be "pure" examples of the clinical problem or diagnosis. In large-scale group treatment studies the problem of multiple or overlapping disorders can be negated by careful subject selection (exclusion criteria) or by random assignment.

For children with complex or multiple problems, there may be one or more other professionals in charge of other aspects of their treatment. In pediatric psychology this is a special problem, as children with chronic illness often have several physicians involved in their care and in addition may be receiving home-based services or allied therapeutic interventions such as physical therapy or nutritional counseling. Treatment provided by other professionals may impact the course of psychological treatment or actually preclude psychological treatment once it has begun. The prescription of psychoactive medications or medications

necessary to treat a medical condition, which have behavioral side effects during the course of psychological interventions, may make it impossible to assess which factor (treatment versus medication) was responsible for noted behavior changes.

In addition to the above difficulties in conducting case studies is the concern over clinical expediency. In group studies, treatment may be financed by a research grant and in some cases patients may even be paid to participate in the research. Most case studies are conducted on patients who are in clinical treatment and are paying for services. Because some procedures that may be required to definitively prove the efficacy of the intervention may extend treatment and therefore increase the cost to the patient, these decisions must be made in consultation with the payer (patient or third-party payer). In the era of managed care, third-party payers are concerned with reducing costs, and parents may not be as concerned as the psychologist about the definitive reason for their child's improvement. Understandably, they are more concerned with effective and efficient treatment. For these reasons, decisions that are scientifically appropriate may not be practically possible. Often a sacrifice must be made between establishing scientific determinations and therapeutic outcome. Patients, parents, and third-party payers are more concerned with efficient outcome than they are with a psychologist publishing a case study. The difficulty in case studies is finding the balance. A proposed treatment plan that is scientifically acceptable may introduce cost or time constraints that result in the patient seeking services elsewhere or in denial of payment by an insurance company. Conversely, treatment that is not well controlled, even if successful, may leave many possible explanations for the reason for the success and may lead to erroneous attributions about treatment components.

TRAINING IN CASE STUDY METHODS

By a wide margin, the number of doctoral-level graduates in clinical psychology prepare themselves for an applied or clinical career. Despite this, most PhD graduate training programs continue to stress group research designs and those with a more behavioral orientation may disproportionately or exclusively train students in single-subject designs. Drotar and colleagues (1995) suggest that the reasons that case studies have become less popular in the *Journal of Pediatric Psychology* relates to the fact that pediatric psychologists have much less experience and training in case study methodologies and that authors and editors do not have comprehensive guidelines for preparing and reviewing case studies. To remedy these concerns, it would seem that graduate-level training in methodologies and considerations appropriate for case study research should receive greater emphasizes.

Earlier in this chapter, concerns about the widening gap between the clinician and researcher was discussed. Remembering that group and single-subject research designs were developed to improve on the case study method of investigation, it is logical that students should be initially introduced to case study methodology so that an appreciation for its potential and for its limitations can form the basis for the understanding of more sophisticated designs. The deemphasis on cases studies and perhaps the disdain of university-based research professors for case study research may contribute to the reluctance of practicing clinicians to think of their cases as

potential research. This needs to change if we are to truly realize the goal of producing research clinicians or clinical researchers. The change must start in our training programs.

On the clinical side, there is a tendency to assume that our interventions are responsible for therapeutic change. An awareness of the multiple factors that may affect behavior, the types of data necessary to document change and, in general, the considerations for evaluating case studies, may serve to make the clinician more objective in evaluating his or her own performance.

SUMMARY

Case reports and case series have played a major role in the development of the fields of clinical child psychology and pediatric psychology. Early case reports served to define the treatment and research parameters of these areas and stand at the crossroads of the scientist and practitioner. While exploratory and descriptive case reports and case series are evident in the literature, it is the explanatory case report that has been the most prominent in the clinical literature. This type of case report attempts to establish cause and effect relationships between an intervention and outcome, generally therapeutic.

Due to lack of scientific rigor and control, case studies are often viewed as less important as they do not provide a basis for drawing scientifically valid conclusions. Rather than being evaluated for providing scientific "proof," case studies are evaluated on their believability. That is, because they lack manipulated variables or scientific controls to ensure that alternative explanations are not operative, case reports must provide information that will allow readers of the reports to logically conclude that other possible explanations for the noted effect(s) are not plausible. Generally, case reports that include objective and numerous sources of data, use frequent measurements, describe the predicted course of the condition or behavior without intervention, and show rapid and large magnitude therapeutic change are more believable. In addition, the more subjects and the less homogeneity in the subjects or conditions, the greater the believability (Kazdin, 1992). A case report is also important in evaluating the author's ability to identify and provide evidence to mitigate or rule out other potential explanations for the noted therapeutic change. Follow-up data also can serve to enhance or diminish the believability of a case report depending on the nature of the treatment and the expected effects.

REFERENCES

- Barlow, D. H., & Hersen, M. (1985). *Single case experimental designs: Strategies for studying behavior change* (2nd ed.). New York: Pergamon.
- Barnett, D. W. (1983). Modified dry bed training: A review, critique and case study. *Journal of Clinical Child Psychology*, 12, 187-191.
- Berkman, K. A., & Meyer, L. H. (1988). Alternative strategies and multiple outcome remediation of severe self-injury: Going "all out" nonaversively. *Association for Persons with Severe Handicaps*, 13, 76-88.
- Busk, P. L., & Marascuilo, L. A. (1992). Statistical analysis in single case research: Issues, procedures, and recommendations, with applications to multiple behaviors. In T. R. Kratochwill & J. R. Levin (Eds.), *Single case research design and analysis: New directions for psychology and education* (pp. 159-186). Hillsdale, NJ: Lawrence Erlbaum.

- Coolidge, C. B. (1977). The dying of Robin: A mother's reflection. *Journal of Pediatric Psychology*, 2, 79–81.
- Cunningham, C. E., & Linscheid, T. R. (1976). Elimination of chronic infant ruminating by electric shock. *Behavior Therapy*, 7, 321–234.
- Dash, J. (1981). Rapid hypno-behavioral treatment of needle phobia in a five-year-old cardiac patient. *Journal of Pediatric Psychology*, 6, 37–42.
- Drotar, D., La Greca, A. M., Lemanek, K., & Kazak, A. (1995). Case reports in pediatric psychology: Uses and guidelines for authors and reviewers. *Journal of Pediatric Psychology*, 20, 549–565.
- Fox, R. M., Bittle, R. G., & Faw, C. D. (1989). A maintenance strategy for discontinuing aversive procedures: A 52 month follow-up of the treatment of aggression. *American Journal of Mental Retardation*, 94, 27–36.
- Hayes, S. C. (1983). The role of the individual case in production and consumption of clinical knowledge. In M. Hersen, A. E. Kazdin, & A. S. Bellack (Eds.), *The clinical psychology handbook* (pp. 181–195). New York: Pergamon.
- Jones, E. E. (1993). Introduction to special section: Single-case research in psychotherapy. *Journal of Consulting and Clinical Psychology*, 61, 371–372.
- Kanner, L. (1957). *Child psychiatry* (3rd ed.). Springfield, IL: C.C. Thomas.
- Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validity. *Behavior Modification*, 1, 427–452.
- Kazdin, A. E. (1992). Drawing valid inferences from case studies. In A.E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 475–490). Washington DC: American Psychological Association.
- Lang, P. J., & Melamed, B. G. (1969). Case report: Avoidance conditioning therapy of an infant with chronic rumitative vomiting. *Journal of Abnormal Psychology*, 74, 1–8.
- Linscheid, T. R., & Landau, R. J. (1993). Going "all out" pharmacologically? A re-examination of Berkman and Meyer's "Alternative strategies and multiple outcomes in remediation of severe self-injury: Going 'all out' nonaversively". *Mental Retardation*, 31, 1–6.
- Linscheid, T. R., Sheldon, K., Loss, N., Ott, J., Mi, J. M., Tschannen, T. (1996). Rapid and effective inpatient behavioral treatment of choking phobia in children. Poster presented at the Society of Pediatric Psychology, Great Lakes Regional Meeting, Toronto, Canada, April 1996.
- Mansheim, P. (1979). Emotional and behavioral data in a case of the 48,XXYY syndrome. *Journal of Pediatric Psychology*, 4, 389–402.
- Noll, R. B., & Seagull, A. E. (1982). Beyond informed consent: Ethical and philosophical considerations in using behavior modification of play therapy in the treatment of enuresis. *Journal of Clinical Child Psychology*, 11, 44–49.
- Ramsden, R., Friedman, B., & Williamson, D. A. (1983). Treatment of childhood headache reports with contingency management procedures. *Journal of Clinical Child Psychology*, 12, 202–206.
- Roberts, M. C. (1992). Vale dictum: An editor's view of the field of pediatric psychology and its journal. *Journal of Pediatric Psychology*, 17, 785–805.
- Sassaman, E. A., Zartler, A. S., & Mulick, J. A. (1981). Cognitive functioning in two sisters with carbamyl phosphate synthetase I deficiency. *Journal of Pediatric Psychology*, 6, 171–176.
- Teichman, Y., & Eliahu, D. (1986). A combination of structural family therapy and behavior techniques in treating a patient with two tics. *Journal of Clinical Child Psychology*, 15, 311–316.
- Tittler, B. I., & Cook, V. J. (1981). Relationships among family, school, and clinic: Toward a systems approach. *Journal of Clinical Child Psychology*, 10, 184–187.
- Van Duser, N., & Phelan, S. (1993). Aversives and the controversy: A view from the parents' perspective. *Child and Adolescent Mental Health Care*, 3, 77–93.
- Walker, C. E. (1995). Elimination disorders. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 537–557). New York: Plenum Press.
- Watson, J., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, 3, 1–14.
- Wells, K. (1987). Scientific issues in the conduct of case studies. *Journal of Child Psychology and Psychiatry*, 28, 783–790.
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.
- Yin, R. K. (1993). *Applications of case study research*. Newbury Park, CA: Sage.

21

Conducting Randomized Treatment Studies in Real-World Settings

SCOTT W. HENGGELER and JEFF RANDALL

With the recent observation by Weisz and his colleagues that efficacious research-based child psychotherapies are not translating to effective community-based practices (Weisz, Donenberg, Han, & Kauneckis, 1995a; Weisz, Donenberg, Han, & Weiss, 1995b; Weisz & Weiss, 1993), the importance of building capacity to conduct rigorous outcome research in community settings has become a priority (e.g., Hoagwood, 1997; Hoagwood, Hibbs, Brent, & Jensen, 1995). The primary purpose of this chapter therefore is to describe strategies that have been successfully used to conduct controlled outcome studies in real-world settings: community mental health centers, private provider entities, and several state agencies. Experiences in conducting clinical trials with multisystemic therapy (MST) (Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998) provides an example of how the rigors of laboratory-based treatment research can be translated into the pragmatics of community-based outcome research or clinical services research (Henggeler, Schoenwald, & Pickrel, 1995; Schoenwald & Henggeler, in press).

In developing community-based outcome studies, in contrast with laboratory-based outcome research, the investigator strives to maximize external validity while minimizing compromises to internal validity (Cook & Campbell, 1979). Thus, with regard to sample selection, exclusion criteria are minimized so that the sample is representative of the intended subset of youths: outpatient mental health cases, violent offenders, children taken into custody for maltreatment, and so forth. Such an approach contrasts with university-based treatment studies that often focus on

SCOTT W. HENGGELER and JEFF RANDALL • Family Services Research Center, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, South Carolina 29425.

highly circumscribed samples, for example, excluding youth from single-parent families or with psychiatric comorbidity, intelligence test scores below 80, medical problems, or serious clinical difficulties. In addition, the community-based study might strive to use practitioners who reflect the demographics, skills, and experiences of clinicians who work in real-world settings. Again, such an approach contrasts with the characteristics of clinicians who typically participate in university-based studies (e.g., graduate students, postdoctorals). The intent is to conduct a study that is high in ecological validity (Bronfenbrenner, 1979) so that findings will have applicability to real-world settings.

On the other hand, to promote the internal validity of findings, community-based outcome studies must endeavor to import the methodological rigor that characterizes university-based treatment research. Such rigor often includes random assignment to treatment conditions, strong treatment specification, significant attention to treatment integrity, and the use of well-validated measurement methods. When the experimental rigor that characterizes laboratory-based research is used within a context high in external validity, the social validity of the findings can be substantial. Indeed, the combination of important results, high external validity, and high internal validity can have a substantive impact on decision makers (Backer, Liberman, & Kuehnel, 1986; Weiss & Weiss, 1981), leading to the dissemination and adoption of innovative mental health services.

Drawing on experiences conducting randomized trials of MST during the past decade, this chapter addresses several issues that are particularly important for investigators interested in conducting well-controlled outcome studies in community settings. In particular, we focus on those aspects of internal validity often neglected by field researchers (e.g., randomization, treatment specification, treatment integrity) and those aspects of external validity often neglected by university-based researchers (e.g., sample generalization, research attrition, treatment retention, interagency collaborations). Methodological issues that have a broad consensus (e.g., using reliable and valid measurement instruments, the importance of follow-up) are not addressed. Finally, ethical issues that are especially pertinent to field researchers are discussed.

PROMOTING INTERNAL VALIDITY

Field researchers of mental health services have tended to neglect several factors that are crucial to the internal validity of findings—the likelihood that observed treatment effects were caused by the treatment (or that treatment had no effects).

Random Assignment to Treatment Conditions

A well-implemented (i.e., high recruitment, low attrition, high treatment integrity, valid measures, etc.) randomized trial provides extremely important and valid information regarding treatment effectiveness. Random assignment controls for confounding factors, such as spontaneous reform (e.g., maturation effects), cohort effects (e.g., decreased crime due to improved economic conditions), and other concurrent interventions (e.g., probation) that may account for outcomes (Henggeler, Smith, & Schoenwald, 1994). To truly establish the effectiveness of a

program, investigators must demonstrate treatment effectiveness beyond what might be expected from confounding factors. Random assignment is the method of choice to ensure that favorable outcomes are due to treatment effects and the absence of random assignment severely compromises the strength of causal inference.

In light of the unequivocal advantages of well-conducted randomized trials in evaluating the effectiveness of a treatment model, we have endeavored to conduct randomized trials in virtually all evaluations of MST. These include: (1) three randomized trials with violent and chronic juvenile offenders (Borduin et al., 1995; Henggeler, Melton, Brondino, Scherer, & Hanley, 1997; Henggeler, Melton, & Smith, 1992); (2) the only published randomized trial with adolescent sexual offenders (Borduin, Henggeler, Blaske, & Stein, 1990); (3) a randomized trial with maltreating families (Brunk, Henggeler, & Whelan, 1987); (4) one of the few randomized trials in the field of children's mental health in which psychiatric hospitalization is one of the treatment conditions (i.e., MST as an alternative to emergency psychiatric hospitalization) (Henggeler et al., in press); and (5) a randomized trial with drug abusing and dependent juvenile offenders (Henggeler, Pickrel, & Brondino, 1999). Each of these studies required extensive partnership with state and community agencies including mental health, juvenile justice, social welfare, education, and the family court. In addition, we are collaborating with academics, stakeholders, and public agencies in ongoing randomized trials of MST in several sites, including Delaware, Texas, Ohio, and Canada. As this list of studies suggests and contrary to prevailing perceptions, rigorous experimental procedures can be implemented in field settings. Gaining the collaboration and support of stakeholders, however, is essential.

Addressing Common Arguments against Randomization

A frequent concern raised by opponents of randomized studies of innovative treatment service models is that some youths will be deprived of the benefit that intensive current services such as residential treatment provide (Henggeler et al., 1994). In such a case, the investigator can argue that providing current services that have little empirical support raises serious ethical concerns (McFall, 1991). What evidence, for example, supports sending youths to residential treatment programs for 12 months at \$60,000 per year. Might there be a better way to spend that money, perhaps on building the capacity of parents to raise their children effectively? Whether the proposed innovative treatment is actually any more effective than current services, however, remains to be seen. The fairest and most valid way to explore these concerns, therefore, is through a well-implemented randomized study. Although this argument will not always convince mid-level staff, funders, and family members will resonate to the issues.

When the innovative treatment is viewed as desirable by stakeholders, a second type of debate against controlled research often ensues. Here, gatekeepers would often prefer to decide which treatment is to be received by the youth and family, as they wish to ensure that the youth receives needed services. The investigator, on the other hand, can argue that under conditions of limited resources (e.g., resources for the pilot project are not sufficient to provide all eligible youths with the innovative treatment) the fairest and most ethical way of allocating scarce resources is through random assignment of youths who meet eligibility criteria.

Otherwise, the unintended biases of gatekeepers will come into play, and these biases might reflect ethnic, socioeconomic, or gender differences.

Gaining Collaboration from Funders of Mental Health Services

In general, funders of mental health services (e.g., state department of health and human services, legislators, policy makers) want to gain the most "bang for their buck." That is, individuals who have fiscal responsibility for meeting the mental health needs of a population are motivated to provide services of demonstrated clinical and cost-effectiveness, services that improve the lives of children while not bankrupting the treasury. Often, these decision makers are dissatisfied with their current portfolio of services and believe that children and families are receiving little benefit from the considerable resources that are being expended. Unfortunately, and as noted above, the research literature tends to provide little support for the clinical or cost effectiveness of existing mental health services.

The failure of current services to document favorable outcomes combined with the high cost of these services (i.e., the majority of funding is devoted to expensive out-of-home placements) provides an opportunity for investigators to argue for the development and rigorous evaluation of innovative mental health services, services designed to achieve favorable outcomes and cost savings. As most policy makers know (Weiss & Weiss, 1981), well-controlled research that includes clinical and cost-related outcomes can provide valuable information regarding the viability of innovations. Consequently, a central argument that we have used to gain the collaboration of funders in conducting randomized trials is that the results will inform policy on two critical fronts: outcomes and costs. Such a perspective usually appeals directly to the needs of funders as well as to the experiences of family advocates, who also are concerned primarily with outcomes and spending resources in the most efficient way possible.

Gaining Collaboration from Administrators and Clinicians

Most of the resistance to conducting randomized trials that examine existing mental health services comes from middle-level managers and practitioners, for whom an evaluation of clinical and cost-related outcomes can be anxiety provoking. Few individuals desire to have their work scrutinized, and accountability for outcomes has traditionally had little place in the field of mental health. Historically, upon receiving their clinical degrees, practitioners have been free to practice as they please, with virtually no examination of the effectiveness of their interventions or the validity of their underlying theoretical and clinical assumptions. Controlled outcome studies that are strongly supported by funders can be viewed as threatening to a comfortable status quo.

Most important for gaining the collaboration of administrators and practitioners involved in the "innovative" treatment is to provide them with a genuine stake in the project. Such a stake can be facilitated in several ways. First, for example, MST training emphasizes that the new treatment model integrates and takes advantage of many of the skills that the practitioners have developed during their careers (e.g., family engagement skills, proficiency in cognitive behavior therapy, behavior therapy, and family therapy; knowledge of indigenous and for-

mal community resources). Second, in several MST project sites the practitioners view themselves as in the vanguard of systemwide change that is forthcoming. Their early support of such change will place them in a relatively favorable position in future years. Third, in several MST sites financial incentives are provided for therapists who adhere to the treatment protocol and obtain favorable outcomes. Fourth, many administrators and practitioners recognize that current services are often not effective with children and families and as such truly wish to help move the field in different directions.

Although the aforementioned strategies have been used effectively to gain the collaboration of stakeholders to conduct randomized trials, in many cases the obstacles to such collaboration are insurmountable. For example, the political and financial incentives of key stakeholders may be contrary to the goals of the randomized trial. Fortunately for investigators, however, mental health services for children are provided across many different service systems (e.g., schools, mental health, juvenile justice, social welfare) and the public and private sector. Thus, multiple opportunities to develop successful collaborations should be available in almost any community.

Treatment Specification

Treatment specification is a critical task in the development, validation, and dissemination of a therapeutic approach. Treatments provided in university settings are often specified through detailed treatment manuals that provide session-by-session guidelines for practitioners to follow. On the other hand, the types of complex, individualized, and comprehensive interventions needed to address the heterogeneous needs of youths and families seen in community settings can be difficult to operationalize and specify (Kazdin, 1988). For example, when working with challenging cases that present serious and diverse crises and problems, MST therapists must be prepared to shift gears at a moment's notice. Moreover, because MST focuses on a wide variety of possible strengths and weaknesses across the social ecologies of youths and families, fully detailing treatment parameters for each possible combination of situations would be an impossible task. Nevertheless, in the absence of strong specification, the value of community-based treatment models is greatly diminished.

To address the problems faced in the specification of MST, a set of nine "treatment principles" has been delineated by the developers of MST to guide therapist behavior (see Henggeler et al., 1998). These principles serve to organize therapists' case conceptualizations, prioritization of interventions, and the types of interventions delivered. For conducting community-based trials, the use of treatment principles has an important advantage over the types of treatment specification that appear in more traditional treatment manuals. Many therapists working in community settings are seasoned professionals who have a wealth of valuable experiences as well as personal strengths and weaknesses. Providing a flexible treatment protocol, within the limits of adhering to treatment principles, allows therapists the freedom to use their strengths to the youth's advantage. In contrast, locking seasoned community-based therapists, many of whom are used to considerable autonomy, into a highly manualized treatment protocol is a recipe for dissatisfaction and staff turnover.

Treatment Integrity

Evaluating and sustaining treatment integrity is no less important in community-based studies than in university-based counterparts. The issue of treatment integrity, however, has historically received little attention in community-based treatment programs.

Promoting Treatment Integrity

To promote treatment integrity in our community-based trials, we have attempted to bring the supervisory intensity common in university-based treatment studies to community settings. Thus, for example, MST therapists in our randomized trials and dissemination projects receive:

1. A week of orientation to the MST treatment model. The objectives of this initial training are to (1) familiarize participants with the scope, correlates, and causes of the serious behavior problems addressed by MST; (2) describe the theoretical and empirical underpinnings of MST; (3) describe the family, peer, school, and individual intervention strategies used in MST; (4) train participants to conceptualize cases and interventions in terms of the principles of MST; and (5) provide participants with role play practice in delivering MST interventions.
2. Quarterly booster training. Sessions are designed to provide training in special topics (e.g., parental substance abuse), address issues that might arise for individuals and agencies using MST (e.g., agency accountability for outcome, interagency collaboration), and allow discussion of particularly difficult cases.
3. Weekly on-site clinical supervision by a senior agency mental health professional trained in MST. Like MST interventions, supervision is pragmatic and goal-oriented. Therapists are expected to conceptualize cases in multi-systemic terms, and supervision is directed toward articulating treatment priorities, obstacles to success, and interventions designed to successfully navigate those obstacles (Schoenwald, Henggeler, Pickrel, & Cunningham, 1996).
4. Weekly consultation via conference call from an expert in MST. This consultation focuses on promoting adherence to MST treatment principles and developing solutions to difficult clinical problems.
5. Ongoing monitoring of therapist adherence to MST principles based on parental reports and through an internet system.

Together, these efforts are intended to promote practitioner adherence to the MST treatment principles and to identify and problem solve any barriers in achieving favorable clinical outcomes for the youth and families.

Therapist Selection

Based on experience in training more than 200 therapists during the past 4 years, Schoenwald and Henggeler (in press) have proposed that several personal and experiential characteristics are positively associated with therapist capacity to implement MST. Important personal qualities include intelligence, flexibility, creativity, and common sense. In addition, therapists with backgrounds in empirically based clinical approaches seem to demonstrate better treatment adherence than

counterparts who have emphasized nonempirically based models (e.g., psychodynamic play therapy, the aesthetic family therapies). Finally, therapists who have volunteered for MST training (versus being conscripted) are receptive to peer supervision and feel accountable for outcome seems to implement MST with greater fidelity.

Measuring Treatment Integrity

Adherence to the MST treatment protocol has been assessed in several ways. First, MST clinical supervisors and administrators have provided ongoing consultation during supervisory sessions, listened to audiorecorded treatment sessions, reviewed case notes, and examined treatment logs to at least qualitatively assess treatment adherence. Second, standardized assessments of treatment fidelity were first obtained by Henggeler et al. (1997) using a 26-item adherence questionnaire that was completed independently by parents, youths, and therapists. As described next, parent ratings of in-session behavior proved to be strong predictors of short- and long-term clinical outcomes. Consequently, the adherence questionnaire has become an important component of the fidelity checks of subsequent MST projects. Third, in an ongoing randomized trial (Henggeler et al., *in press*), we are experimenting with a procedure whereby an MST expert rates 60 minutes of audio-recorded therapist–family interaction each week and provides standardized ratings of therapist MST adherence to the therapist and clinical supervisor the following week. This procedure, though expensive, has improved adherence to the MST treatment principles among the participating therapists.

Empirical Support for the Importance of Treatment Integrity

The value of unwavering attention to treatment integrity was demonstrated in a multisite community-based randomized trial of MST with violent and chronic juvenile offenders (Henggeler et al., 1997). For reasons of convenience, ongoing expert consultation aimed at bolstering treatment fidelity was eliminated from the training and supervisory protocol. At the conclusion of the study, MST outcomes and effect sizes were not as strong as observed in previous trials of MST with like populations. Parent and therapist ratings of adherence to the MST treatment principles, however, predicted long-term outcomes regarding the criminal activity and incarceration of the juveniles. Hence, when treatment fidelity was high, outcomes were favorable, and when fidelity was low, outcomes were weak. These findings have recently been replicated (Henggeler et al., 1999) and have provided valuable lessons regarding the determinants of treatment fidelity.

PROMOTING EXTERNAL VALIDITY

External validity refers to the likelihood that observed treatment effects can be generalized across different types of samples and settings.

Participant Characteristics

University-based treatment research typically focuses on populations of youths who may have little in common with counterparts seen at mental health

centers or at risk for out-of-home placement (Weisz et al., 1995a,b; Weisz & Weiss, 1993). The former are often recruited through advertisements or solicitation from teachers and have circumscribed clinical problems with little comorbidity. On the other hand, children seen in real-world clinical settings tend to have multiple problems (e.g., emotional, behavioral, family-related, school-related) that do not fit into neat diagnostic categories. Moreover, to have resulted in referral to one of the child serving agencies, the problems are usually severe and of long duration.

MST clinical trials typically have sparse exclusion criteria with the intent of both maximizing the generalizability of findings and meeting the needs of collaborators and policy makers. For example, the primary purpose of the Simpsonville Project (Henggeler, Melton, & Smith, 1992) was to test MST as a family- and community-based alternative to the incarceration of violent and chronic juvenile offenders. Inclusion criteria therefore were that the youth (1) was a violent and chronic juvenile offender, (2) was truly at imminent risk of incarceration, and (3) lived with at least one adult family member (i.e., a parent or relative, but not in an out-of-home placement). Youths were not excluded for psychiatric comorbidity, low intelligence test scores, or seriousness of offense history. In fact, youths involved in three deaths were included in the study. On the other hand, due to the desire of the family court in the jurisdiction, youths convicted of selling drugs were sent immediately to detention and not allowed to participate in the project. Similarly, in our study of MST as an alternative to emergency psychiatric hospitalization (Henggeler et al., in press), the goal is to develop a safe, clinically effective, and cost-effective alternative to emergency psychiatric hospitalization. Hence, the study focuses on youths with serious emotional disturbance who have been approved by local mental health authorities for hospitalization due to psychiatric emergencies, and virtually no exclusion criteria pertain to the psychosocial functioning of the youth. Thus, youths with low intellectual functioning, cocaine dependence, psychosis, overt suicidality, and so forth are recruited for participation.

Recruitment of a heterogeneous client sample requires that therapeutic interventions have the breadth to address a wide variety of child and family difficulties. We would argue moreover that all serious clinical populations require comprehensive services provided in the natural ecology (Henggeler & Santos, 1997). Although a particular problem behavior may have drawn the attention of mental health or juvenile justice authorities, the majority of youths presenting with serious clinical problems also evidence difficulties in family, peer, school, and community functioning. From both MST and social ecological perspectives, the goal of treatment is to promote the development of a social ecology that is more conducive to favorable psychosocial functioning for the children and family members. Thus, clinicians must have the capacity to intervene in any area that might significantly attenuate favorable functioning (e.g., parental drug abuse, poor family-school relations) as well as in any area that might promote adaptation (e.g., building indigenous support networks; enhancing parenting skills).

Research Recruitment and Retention Strategies

Modest success at sample recruitment and retention in research can have disastrous consequences for the external and internal validity of findings. For example, if only 60% of eligible families are recruited and only 60% of these

complete the project follow-up, the sample represents only 36% of the identified population. This subset likely differs from the families who refuse to participant and the research dropouts in numerous unknown ways, thus largely invalidating the results of the study. For this reason, MST studies have extended great efforts at participant recruitment and retention. In our trial examining MST with substance abusing and dependent juvenile offenders, 118 of 140 (84%) eligible families were recruited and 97% of these completed the posttreatment assessment (Henggeler et al., 1999). Similarly, in our ongoing study investigating MST as an alternative to emergency psychiatric hospitalization, 90% of the eligible families have been recruited and 99% of these (113 of 114) have completed all assessments thus far.

Although families have been court-ordered into several of our projects, such orders do little to encourage families to cooperate with research protocols, especially follow-up assessments. Rather, several rational strategies can be used to promote research recruitment and retention rates with extremely challenging clinical populations. Foremost is that researchers should appeal to the altruism of the family. That is, family members are informed that the purpose of the research project is to help develop effective treatments for the types of problems presented by their child or adolescent and that the family's cooperation in this endeavor would be of invaluable assistance to others. Second, long-term collaborative relationships should be established with families: family members are fully informed regarding the timing, length, and number of assessments; assessments are scheduled at the family's convenience; contacts are as personalized as possible; families are reimbursed for their participation in assessments (reimbursements have ranged from \$10 per assessment to \$75 per assessment across different projects); the same research assistant follows the family if possible; and researchers behave in friendly and professional manners. Third, each time a family is assessed they should be asked if they are planning to move; regardless of the answer, phone numbers should be obtained of the parent's closest friend, extended family, and the parent's place of work to facilitate tracking. Fourth, research assistants should phone the families after each assessment to thank them and to respond to any negative feedback or complaints that the family members might have. Finally, when a family cannot be located or a phone number has been changed, research staff should immediately attempt to track the family through relatives and friends. For further information, excellent strategies for tracking and retaining participants in longitudinal research projects are provided by Stouthamer-Loeber, van Kammen, and Loeber (1992).

Minimizing Rates of Treatment Dropout

When a treatment outcome study has a high dropout rate, findings are usually uninterpretable. For example, treatment completion rates for substance-abusing youths in therapeutic communities range from 10 to 18% (Pompi, 1994). Even if each treatment completer remained totally abstinent at follow-up, such an outcome could not be validly interpreted. Even if outcomes for the dropouts were known, the completers most likely differ from dropouts in unknown ways and may have ceased drug use no matter what the intervention.

When provided sufficient clinical resources, MST projects have been very effective at engaging families as collaborators in treatment. For example, in our project with substance-abusing and -dependent juvenile offenders (Henggeler et

al., 1999), 98% (57 of 58) of families referred to the MST condition completed a full course of treatment which lasted an average of 130 days (Henggeler, Pickrel, Brondino, & Crouch, 1996). Keys to engagement were (1) a program philosophy that holds the treatment staff accountable for family engagement and outcomes, while providing the resources needed to accomplish these goals; (2) accessibility of services through therapist availability 24 hours per day, 7 days per week, and use of a home-based model of service delivery; (3) a strength-based treatment approach that views family members as full collaborators in treatment: the family takes the lead in setting the treatment goals and the therapist takes the lead in designing interventions to meet the goals; and (4) services are comprehensive and individualized to meet the multiple and changing needs of families.

Building Agency Collaborations

Conducting field-based treatment research requires partnerships between the investigators and community stakeholders. In conducting such a collaboration, all members must benefit and support the value of the project. If any one key stakeholder withdraws from the collaboration, the project can fall apart. For example, the Simpsonville Project (Henggeler et al., 1992) included a collaboration among the Department of Mental Health, Department of Juvenile Justice, and the family court. With families referred from juvenile court (i.e., violent and chronic offenders at imminent risk of incarceration), treatment services provided by a family preservation team located at the community mental health center, and the judges having the authority to do as they wished, any one of these entities could have undermined the study if desired. In recent years, MST projects have included the collaboration of state mental health, juvenile justice, social welfare, education, and health and human services authorities; private agencies that provide mental health services and managed care organizations; private foundations and federal agencies; and judges, public defenders, and prosecutors. Several strategies have been essential in building and maintaining such collaborations.

As described in Henggeler et al. (1998), MST programs build relationships with staff from key agencies by using the types of engagement skills that are successful with families, emphasizing the strengths of the collaborating agency and the common goals between the agency and the MST program. Relations are built and maintained on an individualized basis. For example, in some cases interagency relations are developed and maintained by planning joint social events and gatherings that include all agency members (e.g., a picnic, holiday party). In other cases energies are more efficiently spent by focusing on one or two decision makers within the organization, determining and addressing the needs and concerns of those influential individuals. Moreover, at times we have given fruit baskets to express appreciation for collaborative efforts that go beyond the call of duty or to express an apology for a mistake that we made. Care is taken to maintain lines of interagency communication, and any emerging problems in such communication are addressed immediately, in the early stages. To maintain credibility, the MST program must act with integrity, individuals must keep their word, and interagency "backbiting" must be strictly avoided. In addition, MST programs are encouraged to maintain a healthy quid pro quo, giving to other agencies as much as they receive.

The overriding goal of a partnership is for all collaborators to “win,” youths and families included.

Measuring Outcomes

A crucial aspect of engaging stakeholders in the research process is for the investigators to help the stakeholders define their outcomes of primary interest (e.g., cost per client served, days in out-of-home placement, rearrest) and to build these outcomes into the assessment protocol. Virtually all our field-based randomized trials have been undertaken because key stakeholders desired improved outcomes for children and families, and we were successful in persuading these stakeholders that (1) such outcomes might be attainable if services were restructured in an empirically based fashion, and (2) use of the randomized design would most likely provide valid information regarding these outcomes. Thus, the goals of funders, administrators/practitioners, and families always should be considered important when developing the assessment protocol.

In addition, if funding permits (e.g., in one of our out-of-state randomized trials, funding is only sufficient to examine recidivism and cost; in other trials, \$5 is spent on research for every \$1 spent on services), investigators also should examine those mediating variables that are hypothesized to be linked with the ultimate outcomes of interest. For example, in MST projects that aim to decrease youth criminal activity and drug use, key “instrumental” outcomes include several variables that are both linked with antisocial behavior and a focus of MST efforts. These include, increased parental monitoring of youths, increased family cohesion, decreased youth association with deviant peers, improved youth school performance, and so forth. In light of the ecological nature of MST interventions, measures examine key constructs across the social systems in which the youth is embedded. Because measuring key aspects of the social ecology (e.g., family relations, peer relations, family support network, school performance) can lead to a considerable response burden, we clearly favor those research instruments that are brief and have strong psychometric properties.

ETHICAL ISSUES

MST programs provide ecologically based treatment (e.g., often involving neighbors, school personnel, agency staff, family friends) to address serious clinical problems that often include violence, child maltreatment, and risk of suicide. As such, issues related to confidentiality, risk of violence, and risk of suicide must be constantly scrutinized by MST therapists and supervisors.

Confidentiality

A major goal of MST is to help families address identified problems in their transactions with extrafamilial systems (e.g., peers, school, church, parental workplace, neighborhood, agencies, and so forth) and to build on the existing strengths of the family social support networks. As such, the MST therapist may have direct contact with numerous lay persons and professionals concerned and involved with

the well-being of family members. In exchanging information with these persons, the therapist must have the requisite "release of information" forms signed by the parents. Thus, multiple release of information forms are usually signed during the assessment process so that the therapist can gather information from knowledgeable others. In gaining such information, however, the therapist rarely shares more than very general information about the family with the informant. For example, the therapist might report that Billy has been getting into a lot of trouble with the law and the family is trying to develop strategies to deal with him more effectively; but would not report that mother's cocaine abuse and father's propensity toward violence are interfering with their capacity to parent effectively. Yet, there may be times when carefully targeted disclosure of more extensive information is necessary for family members to reach overarching goals (e.g., to assist mother or father in obtaining services from another agency). When this occurs, the therapist outlines the extent of the information that will need to be disclosed with the family member to obtain his or her approval. In addition, all efforts are made for the family member to be the reporter of sensitive information to, for example, professionals from other agencies as needed.

The issue of confidentiality also pertains to information shared among family members. Drawing on perspectives in the field of family therapy (e.g., Hines & Hare-Mustin, 1978; Margolin, 1982), the MST therapist treats each family member as a separate client with individual confidentiality. However, to increase openness in the family, avoid triangulation of the therapist, and address identified problems more readily, the therapist also informs family members of a "no-secrets" policy. If behaviors reflected in a secret are impacting the functioning of family members, the secret must be addressed by the therapist and secret keeper (if the secret is not impacting family functioning, the secret is not relevant to the situation). The how, by whom, and to whom of addressing the secret will vary with the specifics of the situation, so as to minimize harm and maximize the family's capacity to cope effectively.

Risk of Violence

All mental health professionals are legally and ethically mandated to report ongoing instances of harm to children in order to prevent further harm. Similarly, the therapist is required to report any situation he or she believes may inflict harm on another person. As suggested by Monahan (1993), the therapist should develop a well-specified and documented plan for risk management in conjunction with family members and other professionals. Thus, with regard to issues of safety, the MST therapist has the same ethical and legal mandates as other mental health professionals, and such is noted in the informed consent forms that are completed by families involved in MST projects. When the therapist feels that a report to a social service agency is indicated due to possible harm to a child, he or she first informs the family and requests that they (i.e., the therapist and the parent) contact the appropriate authorities conjointly. The therapist explains why the report must be filed (i.e., protection of the child is primary and it is the law), and what the sequence of events will likely be. If possible, the therapist attempts to attenuate any iatrogenic consequences by appropriately advocating for the family (e.g., emphasizing to authorities that family members are working hard to eliminate risk factors

if such is true). If a therapeutic alliance had been developed before the maltreatment, this straightforward approach may help maintain the therapeutic relationship and provide an opportunity to work with the family on addressing the events that led to the report.

On the other hand, the MST practitioner often treats families that include violent juveniles engaging in criminal offending. The expressed purpose of the MST services is to maintain youth with their families and in the community if at all possible. Importantly, the viability and health benefits of this goal are well documented, even with violent juvenile offenders. For example, in the Simpsonville Project (Henggeler et al., 1992) with violent and chronic juvenile offenders, youths receiving MST in comparison with youths receiving usual services: (1) had significantly fewer arrests at a 59-week follow-up, (2) reported significantly less criminal activity at posttreatment, and (3) were in the community for an average of 73 more days during the 59-week follow-up. Thus, even though approximately 70% of the youths in the usual services condition were incarcerated to promote community safety, these youths still managed to engage in more criminal activity during follow-up than did their counterparts who received MST. In fact, because association with deviant peers is a powerful predictor of criminal activity (Henggeler, 1991), the antisocial behavior of the offenders should be exacerbated by prolonged placement with antisocial peers. Alternatively, by addressing the known causes of criminal activity in an ecologically valid context, MST reduced rates of offending and served to promote community safety. Thus, MST therapists do not function as probation officers (e.g., tracking and reporting criminal behavior), but aim to facilitate parents in restructuring the youth's social network to decrease antisocial behavior and increase prosocial behavior. Nevertheless, MST therapists, as noted above, must maintain ethical and legal standards regarding risk of harm to others (e.g., we have advocated for the incarceration of youths who were an immediate danger to the community).

Harm from Self

Youths at risk of harming self are often hospitalized in psychiatric facilities to protect them. As Sondheimer, Schoenwald, and Rowland (1994) have documented, however, psychiatric hospitalization of children and adolescents has no proven benefits and several significant disadvantages pertaining to cost and the negative effects of hospitalization on youths (e.g., Weithorn, 1988). To address the downside of psychiatric hospitalization as a means to protect the safety of youths at risk of harming themselves, we are currently conducting a randomized trial of MST as an alternative to the hospitalization of youths presenting psychiatric emergencies (Henggeler et al., in press). An important aspect of the clinical protocol in this study includes the development of a highly specified safety plan that uses indigenous resources (e.g., immediate family, extended family, friends, neighbors) to ensure the safety of youths who remain in the community. As per all MST programs, treatment also focuses on ameliorating those factors in the youth's social ecology that were associated with the psychiatric crisis. Thus, by empowering indigenous resources to safely manage crises and by directly addressing problematic aspects of the social ecology so that crises are less likely to occur, MST may pose important long-term advantages over psychiatric hospitalization.

ACKNOWLEDGMENTS. Preparation of this manuscript was supported by the following grants: R01MH51852 from the National Institute of Mental Health; R01DA08029 and R01DA10079 from the National Institute on Drug Abuse, 96.2013 from the Annie E. Casey Foundation, and 97-3N-FX-0016 from the Office of Juvenile Justice and Delinquency Prevention.

REFERENCES

- Backer, T. E., Liberman, R. P., & Kuehnle, T. G. (1986). Dissemination and adoption of innovative psychosocial interventions. *Journal of Consulting and Clinical Psychology*, 54, 111-118.
- Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, 35, 105-114.
- Borduin, C. M., Mann, B. J., Cone, L. T., Henggeler, S. W., Fucci, B. R., Blaske, D. M., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders: Long-term prevention of criminality and violence. *Journal of Consulting and Clinical Psychology*, 63, 569-578.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Brunk, M., Henggeler, S. W., & Whelan, J. P. (1987). A comparison of multisystemic therapy and parent training in the brief treatment of child abuse and neglect. *Journal of Consulting and Clinical Psychology*, 55, 311-318.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton Mifflin.
- Henggeler, S. W. (1991). Multidimensional causal models of delinquent behavior and their implications for treatment. In R. Cohen & A. Siegel (Eds.), *Context and development* (pp. 211-231). Hillsdale, NJ: Lawrence Erlbaum.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology*, 65, 821-833.
- Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology*, 60, 953-961.
- Henggeler, S. W., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance abusing and dependent delinquents: Outcomes, treatment fidelity, and transportability. *Mental Health Services Research*, 1, 171-184.
- Henggeler, S. W., Pickrel, S. G., Brondino, M. J., & Crouch, J. L. (1996). Eliminating (almost) treatment dropout of substance abusing or dependent delinquents through home-based multisystemic therapy. *American Journal of Psychiatry*, 153, 427-428.
- Henggeler, S. W., Rowland, M. R., Randall, J., Ward, D., Pickrel, S. G., Cunningham, P. B., Miller, S. L., Edwards, J. E., Zealberg, J., Hand, L., & Santos, A. B. (in press). Home-based multisystemic therapy as an alternative to the hospitalization of youth in psychiatric crisis: Clinical outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*.
- Henggeler, S. W., & Santos, A. B. (Eds.). (1997). *Innovative services for "difficult to treat" populations*. Washington, DC: American Psychiatric Press.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M.D., & Cunningham, P. B. (1998). *Multisystemic treatment of antisocial behavior in children and adolescents*. New York: Guilford Press.
- Henggeler, S. W., Schoenwald, S. K., & Pickrel, S. G. (1995). Multisystemic therapy: Bridging the gap between university- and community-based treatment. *Journal of Consulting and Clinical Psychology*, 63, 709-717.
- Henggeler, S. W., Smith, B. H., & Schoenwald, S. K. (1994). Key theoretical and methodological issues in conducting treatment research in the juvenile justice system. *Journal of Clinical Child Psychology*, 23, 143-150.
- Hines, P. M., & Hare-Mustin, R. T. (1978, February). Ethical concerns in family therapy. *Professional Psychology*, 165-171.

- Hoagwood, K. (1997). Interpreting nullity: The Fort Bragg Experiment—A comparative success or failure? *American Psychologist*, 52, 546–550.
- Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. (1995). Introduction to the special section: Efficacy and effectiveness in studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63, 683–687.
- Kazdin, A. E. (1988). *Child psychotherapy: Developing and identifying effective treatments*. New York: Pergamon.
- Margolin, G. (1982). Ethical and legal considerations in marital and family therapy. *American Psychologist*, 37, 788–801.
- McFall, R. M. (1991). Manifesto for a science of clinical psychology. *Clinical Psychologist*, 44, 75–88.
- Monahan, J. (1993). Limiting therapist exposure to Tarasoff liability. *American Psychologist*, 48, 242–250.
- Pompi, K. F. (1994). Adolescents in therapeutic communities: Retention and posttreatment outcomes. In F. T. Tims, G. De Leon, & N. Jainchill (Eds.), *Therapeutic community: Advances in research and application* (pp. 128–161). National Institute on Drug Abuse Monograph 144. Rockville, MD: National Institutes of Health Publication No. 94-3633.
- Schoenwald, S. K., & Henggeler, S. W. (in press). Services research and family based treatment. In H. Liddle, G. Diamond, R. Levant, & J. Bray (Eds.), *Family psychology intervention science*. Washington, DC: American Psychological Association.
- Schoenwald, S. K., Henggeler, S. W., Pickrel, S. G., & Cunningham, P. B. (1996). Treating seriously troubled youths and families in their contexts: Multisystemic therapy. In M. C. Roberts (Ed.), *Model programs in child and family mental health* (pp. 317–332). Mahwah, NJ: Lawrence Erlbaum.
- Sondheimer, D. L., Schoenwald, S. K., & Rowland, M. D. (1994). Alternatives to the hospitalization of youth with a serious emotional disturbance. *Journal of Clinical Child Psychology*, 23(Suppl.), 7–12.
- Stouthamer-Loeber, M., van Kammen, W., & Loeber, R. (1992). The nuts and bolts of implementing large-scale longitudinal studies. *Violence and Victims*, 7, 63–78.
- Weiss, J. A., & Weiss, C. H. (1981). Social scientists and decision makers look at the usefulness of mental health research. *American Psychologist*, 36, 837–847.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Kauneckis, D. (1995a). Child and adolescent psychotherapy outcomes in experiments versus clinics: Why the disparity? *Journal of Abnormal Child Psychology*, 23, 83–106.
- Weisz, J. R., Donenberg, G. R., Han, S. S., & Weiss, B. (1995b). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63, 688–701.
- Weisz, J. R., & Weiss, B. (1993). *Effects of psychotherapy with children and adolescents*. New York: Sage.
- Weithorn, L. (1988). Mental hospitalization of troublesome youth: An analysis of skyrocketing admission rates. *Stanford Law Review*, 40, 663–838.

22

Evaluating Mental Health Services for Children and Adolescents

**SUSAN DOUGLAS KELLEY, CAROL T. NIXON,
and LEONARD BICKMAN**

Aside from a different emphasis on intervention, such as job training or housing for adults, many have characterized a split between child and adult mental health services research (e.g., Mechanic, 1996). The research, theory, knowledge, and legislation in children's mental health services has lagged far behind that in the adult field (Institute of Medicine, 1989). This chapter provides an overview of key issues in children's mental health services research. First, the definition and functions of mental health services research are discussed. Although the field is still developing, mental health services research serves several functions, primarily to provide an objective means of informing public policy and program decision making, and ultimately contributing to the improvement of clinical practice. A distinction is then made between efficacy research, or treatment research, and effectiveness research, which also can be called services research. The discussion focuses not only on the differences between the two, but also suggests the importance of the complementary use of both approaches in order to increase our knowledge of "what works" in children's mental health services in applied settings.

Conducting services research in the "real world" is difficult, and relevant methodological and practical issues are rarely addressed in training programs. Consequently, there is a need for the description of methods and management of methodological dilemmas. This chapter focuses on approaches necessary to strengthen research in children's mental services, from the use of program theory in defining evaluation studies to the importance of building collaborative relation-

SUSAN DOUGLAS KELLEY, CAROL T. NIXON, and LEONARD BICKMAN • Center for Mental Health Policy, Vanderbilt University, Nashville, Tennessee 37212.

ships among stakeholders, including researchers, policy makers, administrators, clinicians, and the children and families being served. Recent evaluation studies will be included to describe the lessons learned and innovative approaches to common obstacles encountered in the field.

The field of children's mental health services currently finds itself in a state of confusion. Most effectiveness studies have not demonstrated the effectiveness of children's mental health programs. This has led to the recognition that more high-quality research on what works in real-world settings is greatly needed. Moreover, such null findings from well-designed studies point to the possibility that the basic assumptions in the field of children's mental health services need to be revisited. Without being able to adequately describe treatment, the evaluation of programs or whole systems of care may continue to show little effect on client-level outcomes. Furthermore, the clinical relevance of services research can be emphasized by increasing the active participation of direct service providers in the research process. These issues are discussed in the last two sections of the chapter, followed by the implications for the current state of the field and suggestions for future research.

WHAT IS MENTAL HEALTH SERVICES RESEARCH?

Mental health services research is frequently defined as consisting "mainly of those studies examining the *effectiveness* [see discussion to follow] of the health care system in the prevention, diagnosis, and treatment of individuals with or at risk of mental disorders" (Kelleher & Long, 1994, p. 133). Starfield (1973) classically delineated areas within services research, distinguishing between clinical (or process) and systems (or structural) research. Clinical health services research is narrower in focus including, for example, investigation of the outcome of specific interventions and the relationship between clinicians and clients. Systems research is typically broader, encompassing organizational and financial factors that affect the delivery or quality of health care (Kelleher & Long, 1994).

Mental health services research is still in its infancy and basic issues about services are still being examined. For example: How much do systems of care cost? How can cost best be controlled? Which models of care lead to the best outcome? How can quality and outcome be defined, assessed, and maintained? How does interagency coordination affect the delivery and quality of services? How and to what extent should multiple stakeholders contribute and participate in the development and delivery of services? Due to the complexity of all the features that contribute to a quality mental health service system and in light of system reform and managed care, more emphasis needs to be placed on soliciting the input of multiple stakeholders, including researchers, clinicians, policy makers, and clients and their families. These perspectives can help to clarify the identification of priorities and needs, as well as the actual evaluation of programs and services.

FUNCTIONS OF MENTAL HEALTH SERVICES RESEARCH

The primary goal of mental health services research is to inform policy and programmatic decision making in order to improve and enhance the quality of

mental health services. Politically, the evaluation of programs, services, and systems provides a relatively objective means of choosing among alternative programs or services competing for limited resources and informing decisions that will affect public policy, both directly and indirectly. With the proliferation of managed care in the health services, services research that examines the impact of managed care on service delivery as well as the associated outcome will help to guide public policy. Services research can be simply descriptive, for example, determining who is receiving what types of services at what cost. However, it also can be used to inform public policy by examining the impact of external constraints (e.g., health insurance companies, entitlement programs, and social welfare legislation) on individual-, service-, and system-level outcomes. For the purposes of this chapter, the focus is on the evaluation of services, namely program process and outcomes.

Evaluation in general encourages accountability, cost consciousness, and responsiveness to those in need of and utilizing services both at the system and program levels, that is, those at the policy level as well as those actually delivering services. Indeed, a recent emphasis has focused on outcomes of services and systems of care. Efforts to develop and utilize performance indicators and outcome measurement have increased. Services research can provide information about the services that are actually delivered as compared to those intended, how services are being used and by whom, the extent to which services are effective, and the costs of services. In fact, Bickman (1992) has pointed out it may be unethical to continue to provide untested, perhaps ineffective services to children and their families in the name of "treatment."

Further, services research can synthesize what is already known, unearth false assumptions, debunk myths, develop new information, and explain the implications of this information for future decision making, again at both the system and program levels (Chelimsky, 1995). In order to make such an impact, Friedman (1997) stressed that numerous evaluations over time using multiple methods will help to build a solid base of knowledge about providing mental health services to children and their families. The existing literature not only provides a means of identifying existing gaps in the field, suggesting topics in need of study, but also provides the foundation for analytic strategies such as meta-analysis, which is useful for synthesizing the existing research. Ultimately, services research not only contributes to political decision making but also aims to improve clinical practice. Multidisciplinary research guided by the input of multiple stakeholders, including the clinicians and their clients, in all stages of research (conceptualization, research design, implementation, monitoring, dissemination of findings, etc.) is designed to increase the relevance, fairness, and utility of services evaluations (Weiss, 1983a,b). For example, clinicians and services researchers jointly can examine patient selection processes, effectiveness of interventions, impact of financial and organizational factors on selection of treatment settings, choice of treatment, intensity of care, and provider mix (Mechanic et al., 1992).

THE CRITICAL DISTINCTION BETWEEN EFFICACY AND EFFECTIVENESS

As noted previously, one aim of mental health services research is to examine the "effectiveness" of services or systems of care. In the health services research field, a basic distinction has developed between the efficacy of an intervention and

its effectiveness. Studies of the efficacy of an intervention examine the relationship between intervention processes and outcomes under tightly controlled or optimal conditions. The investigator usually has substantial control over sample selection, the intervention, its delivery mechanisms, and the environmental setting. For example, randomized clinical trials (RCTs), often conducted within a university-based laboratory, are typically efficacy studies.

In contrast, studies of effectiveness examine the impact of interventions in naturalistic settings or under typical conditions of use. Studies of effectiveness usually are conducted in real-world clinical settings [i.e., schools, general medical settings, community mental health centers, or health maintenance organizations (HMOs)] where the researcher loses much of the control over factors such as sample selection, fidelity of services delivery, and choice of clinicians. Effectiveness should be defined broadly as allowing examination of such issues as service access, utilization, cost, and quality in addition to issues of service impact (outcome evaluation). However, in comparison to treatment research, services research deemphasizes specific disorders or conditions. Because applied research has different objectives than basic research (Bickman, Rog, & Hedrick, 1997b), it is not surprising that these differences exist.

Effectiveness research is a key goal of services research, although efficacy and effectiveness research are both important. Hoagwood, Hibbs, Brent, and Jensen (1995) outlined a five-phase model developed by the National Cancer Institute that serves as a basic, yet ideal, framework for conducting research. The steps are the following: (1) hypothesis development; (2) methods/instrumentation development and refinement; (3) controlled intervention trials such as RCTs; (4) defined population studies in which subgroups of the population are studied; and (5) demonstration and implementation studies that examine the transfer of efficacious interventions to more applied settings. The latter are studies of effectiveness.

A plethora of efficacy studies of psychological treatments and services (primarily psychotherapy) have been conducted in both the adult and children's mental health fields as have several major meta-analyses (e.g., Kazdin, Bass, Ayers, & Rodgers, 1990; Lipsey & Wilson, 1993; Lonigan, Elbert, & Johnson, 1998; Smith & Glass, 1977; Weisz & Jensen, 1999; Weisz, Weiss, Aliche, & Klotz, 1987). Although many of the individual studies have failed to demonstrate significant effects, meta-analytic reviews have produced "stark, dramatic patterns of evidence for the general efficacy of such [mental health] treatment" (Lipsey & Wilson, 1993, p. 1182).

Relative to the number of efficacy studies, few studies have examined the effectiveness of mental health interventions or services, particularly in regard to children (Weisz, Weiss, & Donenberg, 1992). More investigations of this type are greatly needed (National Institute of Mental Health, 1990). The effectiveness research that has been conducted has focused primarily on system-level variables (e.g., administrative and financing structures, service coordination, access, utilization) with little investigation of clinical outcome (Friedman, 1997; Henggeler, Schoenwald, Pickrel, Rowland, & Santos, 1994; Speer & Newman, 1996) and specific mental health treatments (Farmer, Burns, Guiles, Behar, & Gerber, 1997; Institute of Medicine, 1989). According to Henggeler et al. (1994), "the dearth of clinical outcome data in the services research literature raises the frightening possibility that evidence of increased access and variety of services may be construed as a proxy for quality and effectiveness of clinical services rendered" (p. 230).

In fact, most of the studies that have examined the effectiveness of mental health services have not demonstrated a general link between the intervention studied and improved clinical outcomes for children or for adults (Speer, 1994; Weisz, Catron, Harris, & Phung, 1999; Weisz et al., 1992). For example, the Fort Bragg Evaluation Project (FBEP) found that children in a continuum of care had increased access to the appropriate services compared to children at the control sites who only received traditional inpatient and outpatient services. Both the experimental and control groups improved but equally as much; the clinical and family outcomes did not differ between the two groups (Bickman et al., 1995a; Bickman, 1996a,c, 1997). Further, a multisite Robert Wood Johnson Foundation project did not find a relationship between system changes and clinical outcome for the seriously mentally ill adults who were served (Morrissey et al., 1994). However, other researchers have obtained some promising results demonstrating the potential effectiveness and cost impact of assertive community treatment for adults with severe emotional illness and multisystemic therapy for adolescents with serious emotional disturbance (e.g., Henggeler, Schoenwald, & Pickrel, 1995; Santos, Henggeler, Burns, Arana, & Meisler, 1995).

Several researchers have proposed potential explanations that might account for the differences in findings among efficacy and effectiveness studies, including methodological reasons such as small sample sizes, poor theoretical framework, inappropriate instrumentation, and so forth (e.g., Bickman, 1992; Burns, 1994; Hoagwood, 1994; Kendall & Southam-Gerow, 1995; Lipsey, 1989, 1990; Weisz, Donenberg, Han, & Weiss, 1995). One possible reason for the difference in clinical findings between efficacy and effectiveness studies is the concern that practitioners are not using treatments that have been found to be effective. Because of this strong possibility, professional societies are developing treatment guidelines that reflect the findings from these efficacy studies. However, it is still to be determined whether these techniques can be transferred to the real world and whether practitioners will adopt them. Moreover, it is far from clear that guidelines are specific enough to affect practice.

More attention needs to be devoted to conducting well-designed effectiveness studies in order to bridge the gap between positive outcome efficacy and null outcome effectiveness findings. Research is needed that examines the components or specific processes within an intervention in order to determine those that are the most effective and under what circumstances.

ISSUES IN STRENGTHENING CHILDREN'S MENTAL HEALTH SERVICES RESEARCH

There are critical differences between efficacy and effectiveness research in the trade-offs between internal and external validity and between construct validity and external validity. In the latter situation, efficacy researchers have great control over the intervention and can be more certain that the intervention, as implemented, reflects the theory or construct underlying the intervention. In this type of research the intervention is likely to be "purer" and not confounded with other variables. In effectiveness research the intervention is more likely to be a complex package, and this will make it more difficult to pinpoint the components of the

intervention that were responsible for any effect. On the other hand the effectiveness setting helps ensure that the intervention and results will generalize to real-world settings.

Another possible trade-off is between internal validity and external validity. *Internal validity* is the extent to which a causal linkage can be drawn between the intervention and the desired outcomes by ruling out alternative explanations of the observed effect (e.g., maturation). The tighter the control held by the experimenter, the more likely he or she will be to rule out explanations of observed change other than the studied intervention. On the other hand, *external validity*, as noted earlier, refers to the likelihood that the results observed in one investigation would be repeated, or would generalize, to other settings, samples, or times. These types of validity are extensively discussed by Campbell and colleagues (Campbell & Stanley, 1966; Cook & Campbell, 1979).

A specific example of the trade-offs between design and external and internal validity is the comparison between the method used to recruit clients in the FBEP (Bickman et al., 1995a) with those used in the Stark County Evaluation Project (SCEP) (Bickman, Summerfelt, Firth, & Douglas, 1997c; Bickman, Summerfelt, & Noser, 1997d). The former was a quasi-experiment that allowed the recruitment of severely ill children, since assignment to services was not made by the investigator but was determined simply by where the child lived. Thus, a significant proportion of youth were in the hospital when they were recruited. In contrast, the Stark County project used random assignment with several exclusionary criteria, such as not including emergency cases. Since the children involved in the study were randomly assigned to the demonstration project or to traditional services, it would be unethical to exclude emergency cases from what was thought to be the best care available in the community. Moreover, since the baseline data were collected before the child was assigned to treatment and before the child could receive services, there would be an unacceptable delay in the receipt of services. While the randomized experiment had better internal validity, the quasi-experiment had better external validity. However, the complementary strengths and weaknesses of both designs and the consistent results from both studies make the similar findings more credible.

Real-world delivery of interventions tends to include heterogeneity in the client population, the services, and the environment in which they are delivered. According to Cronbach et al. (1980), evaluation has placed too high an emphasis on internal validity. Threats to internal validity always are present, and attempts to control them often lead to the evaluation of the intervention in an artificial setting, which is unlike the setting in which interventions are actually delivered. In order to carry out such evaluation studies in the "real world," researchers need to attend to practical strategies that will serve to enhance internal validity while also preserving the purpose of evaluating actual programs. Several recent articles have specified threats to internal validity and other problems in applied research (e.g., Bickman, 1992; Bickman, Summerfelt, & Foster, 1996; Burns, 1994; Hoagwood, 1994; Lipsey, 1990; Persons, 1991; Bickman & Rog, 1997). Although an exhaustive summary will not be repeated here, this section will describe some of the issues to be resolved in planning evaluation studies in children's mental health services research. Several recent evaluation studies will be used to illustrate some of the common problems that need to be considered, providing an idea of the innovation that has advanced the research methodology in the area of children's mental health services research.

Program Theory

Bickman (1987) defined program theory as “a plausible and sensible model of how a program is supposed to work” (p. 5). Although it has been suggested that all social programs are based on a general theory of how or why that particular program should work, much of the time such theories are not explicitly stated (Weiss, 1995). In order to test the linkages between program components and planned outcomes, explication of the nature of the program itself is an important first step in any evaluation. Researchers should identify assumptions and theories of change held by program staff and other key stakeholders in order to assess these in the process of evaluation. Furthermore, efforts need to be devoted to the development of conceptual models that allow exploration of particular components of care for specific segments of the target population. *What* is effective, for *whom*, and under what *circumstances*? Without thoroughly defining the logic underlying the program, thus enabling a more comprehensive and targeted assessment, it is impossible to determine whether a lack of positive results is due to failure of the program theory itself or failure to accurately test the program (program implementation failure).

Despite this basic understanding of the need for program theory, evaluation has been criticized as atheoretical and in dire need of attention to theory at the outset of research efforts (e.g., Bickman, 1987; Boruch & Gomez, 1977; Chen & Rossi, 1983; Weiss, 1983a,b, 1995). Although the field of children’s mental health services has been successful in disseminating guidelines for care thought to impact on mental health services delivery, “there has been much talk about values and concepts but very little talk about theory” (Friedman, 1997, p. 25). For example, the Child and Adolescent Service System Program (CASSP) was developed by the National Institute of Mental Health to support state and local systems of care for children and youth with serious emotional problems and their families (Stroul & Friedman, 1986, 1996). Based on findings from innovative systemic intervention models from the early 1980s, the CASSP developed principles to be followed in the creation of systems of care for children, including that services should be family focused, children should be served in the least restrictive environment, services should be appropriate to the child and family’s needs, and services should be community-based. Although these efforts have had a major influence on the structure of children’s mental health services across the country, Friedman (1997) has pointed out that the lack of well-controlled evaluations of such programs limits our ability to interpret findings and provide adequate support to ensure the field is moving in an appropriate direction. Cross and Saxe (1997) echoed this concern, noting that only articulated theories of systems of care in conjunction with outcome research are likely to further our knowledge in hopes of improving the quality of services and systems for children with emotional and behavioral disorders.

Evaluation calls for at least two types of theoretical models (Nixon, 1997). The first represents the development of “general understandings of social phenomena” (Chen & Rossi, 1983, p. 285). It synthesizes the empirical findings in the literature as to how some characteristic, for example, depression, arises and changes over the course of development and identifies interactional variables such as gender. The second type of model, the program theory, pertains to a specific intervention and specifies program resources, activities, and intended outcomes. The first type of theory is nonspecific, applicable to many particular interventions, and most impor-

tantly is imbedded within the program model, linking the pieces—resources, activities, and outcomes—together. It frequently is the “glue” missing in program models, explaining the use of resources and intervention activities in order to achieve desired outcomes. Moreover, it is from theories of social phenomena that program developers and evaluators construct hypotheses about exogenous and intervening variables and how they potentially may affect outcome variables. Finally, according to Weiss (1995), “evaluations that address the theoretical assumptions embedded in programs may have more influence on both policy and popular opinion” (p. 69).

For example, the program theory for the demonstration project in the FBEP (Bickman et al., 1995a; Bickman, 1996b) included several underlying assumptions based on the CASSP principles related to intake, assessment, and treatment of children's mental health problems. According to this model, a quality system of care emphasizes continuity between services, has fewer dropouts from treatment, provides individualized treatment at the most appropriate and least restrictive level, and has more timely transitions between levels of care. These principles were the “glue” hypothesized to explain the link between specific components of the demonstration service system and outcomes such as improved mental health, lower costs per case, quicker recovery, and more client satisfaction. However, Bickman (1997) points out that because the evaluation was a systems-level study, the nature of the treatment itself was studied in broad terms (e.g., length of treatment, type of service placement) but information about specific interventions was not included.

Theory-Based Measurement

Such a comprehensive theoretical model is essential to identifying and operationalizing the constructs to be measured and alerting the evaluators to the presence of untested assumptions. This includes determining the level of analysis at which the study is designed and developing assessment strategies to measure both program operations and program-related outcomes relevant to the theory underlying the program. For example, the Fort Bragg Evaluation Project (Bickman et al., 1995a; Bickman, 1996b) was designed to examine the impact of a continuum of care on children's mental health outcomes and cost of services. Thus, the level of analysis for the study was focused on system change, such as the introduction of prompt intake, comprehensive multidisciplinary assessment, family participation, and case management, not the evaluation of services such as psychotherapy or hospitalization or processes such as assessment or treatment planning. Furthermore, Bickman (Bickman, 1996b; Bickman & Salzer, 1997) suggests that no standards currently exist in the field of children's mental health to adequately measure the relative quality of such aspects of care as assessment or treatment plans. While the program theory was informative in indicating links that would remain untested in the FBEP, it also provided the stimulus for future research into the appropriateness and quality of treatment (Bickman, 1995, 1997).

In order to measure children's mental health services, it is necessary for researchers to develop both clear definitions of the program-related activities of interest and systematic data collection procedures. Because most evaluations focus on the frequency and types of mental health services utilized, important components of information gathering may include noting the timing of services, the cost of

services to either the family or the program, and who was involved in delivering services to children with emotional or behavior problems and their families. To collect such information, the researcher can modify the existing record keeping done by program staff and practitioners or develop new strategies for collecting information, such as incorporating a management information system (MIS) into the structure of the program itself. Abstracting information from case records is another method available, but it is usually more costly and time-consuming. Furthermore, if there is no prior agreement on the specific information to be noted in the records, needed information may not be available. Therefore, collaborating with program staff and clinicians to clearly define the service units to be measured and incorporating evaluation instruments into ongoing record-keeping activities can serve to improve the quality and usefulness of the data collection for evaluation research.

In addition to assessing program process, the inclusion of clinical outcome data is vital to providing support to the belief that innovative mental health programs can improve children's mental health outcomes. To conduct a comprehensive assessment of children's psychological functioning in a variety of domains, it is important to include multiple informants (such as children, primary caregivers, teachers, therapists, and interviewers) as well as multiple measurement approaches (such as interviews, checklists, and archival records and if possible direct observation). Although structured diagnostic interviews are comprehensive indicators of psychopathology, they are resource-intensive in terms of cost and staffing and they may not be necessary to answer services research questions.

In order to address these concerns, Bickman et al. (1997c) incorporated interactive computerized interviews (Summerfelt, 1992; Summerfelt & Hodges, 1992) into their primary data collection procedures for the SCEP. Completed by both children (over 9 years of age) and their primary caregivers, the computerized interviews were modified from a semistructured interview created to gather information on children's psychiatric symptoms and diagnoses. While a more rigorous assessment of the use of interactive computerized interviews is needed, the authors asserted that the feasibility of using these assessments was demonstrated (Bickman et al., 1997c). Such innovations in measurement have advantages that may serve to reduce the financial burden on funders of research, potentially enabling greater sample sizes and/or more frequent assessments and longer follow-up phases, all needed to further the current state of the field.

Because of the narrower scope of services research, namely, the focus on the overall impact of a service system rather than individual differences in outcomes, it is important to remember that the unit of analysis is not at the level of the child and family. Instead, Friedman (1997) suggested that outcomes related to each child and family serve "as a test of how well the system is implementing the explicitly desired principles and practices" (p. 31). Thus, since one of the goals of services research is to minimize the burden of data collection, multiple informants and multiple measures in the assessment of clinical outcomes may be viewed as secondary to assessing program process, given the limitations in resources. This recommendation is tentative, but some preliminary research appears to indicate that for practical concerns, like predicting service use, in contrast to diagnosis, a parent informant appears sufficient (Bickman, Lambert, Northrup, Salzer, & Summerfelt, 1997a). Moreover, other studies indicate that for producing an aggregate measure of severity of outcome, there is a very high intercorrelation between observers and measures (Lambert, Nixon, Simpkins & Bickman, 1996).

Finally, although the importance of theoretically based assessment methods has been emphasized here, Bickman (1990) has recommended that the “kitchen sink” approach should be used in program evaluation. This approach stresses the importance of improving statistical conclusion validity over construct validity. Here, the experimenter attempts to make the intervention as powerful as possible by throwing in everything, including the kitchen sink, in order to increase the power of the intervention. The trade-off is that while this approach may produce both statistically and clinically meaningful results, the more complex the intervention, the more difficult it is to identify which components of the intervention worked and why. However, a pure but weak intervention will usually not produce significant effects. Moreover, the chances are higher to obtain support to determine which factors were responsible for an intervention in contrast to obtaining support to test another weak intervention.

While maintaining a strong theoretical base, researchers should include measurements of other possible relevant components of the program, particularly with new programs. Factors not accounted for in the program theory should be measured in order to “rule out” alternative explanations. This does not mean developing a complicated research design, which is difficult to implement in the field. Instead, multiple types and sources of data should be included to examine and understand the program process. In addition, decisions regarding the comprehensiveness of assessing both program process and clinical outcomes may be constrained by the resources available to the researchers. However, attempts to rule out alternative explanations are especially relevant to the evaluation of children’s mental health services, given the number of studies that have failed to demonstrate the effectiveness of innovative intervention programs. As Bickman (1990) stated, “Establish the effect first ... if the new program fails to produce an effect there is usually not a second chance” (p. 143).

Implementation Monitoring

Aside from program theory failure, program evaluators must address problems of program implementation failure (Bickman, 1987, 1990; Chen & Rossi, 1992; Lipsey, 1989, 1990). A number of evaluators have emphasized the importance of monitoring treatment over the course of the evaluation in order to ensure that the planned treatment did in fact occur (e.g., Burns, 1994; Chen, 1990; Heflinger, 1996a; Rezmovic, 1984; Scheirer, 1987). Detailing the “black box” by including information about how a program is put in place is critical to understanding what the program accomplishes. Without ensuring that the treatment did occur and specifying individual components and activities of the intervention, it is difficult to draw useful conclusions about the program (Heflinger, 1996a). If the program succeeds, there is not enough information to replicate it in the future. If the program fails, it is not clear whether the apparent failure was due to the program itself or to problems with the implementation of the program. Rezmovic (1984) stated that

it is impossible to know how often studies have concluded that treatments were ineffective when they should have concluded that shortcomings in treatment implementation precluded the drawing of conclusions about their value” (p. 189).

The cumulative effect of policy decisions based on such studies can have serious effects on the future of children's mental health services.

Therefore, while planning an evaluation study, researchers should allocate funding resources to conduct an implementation study and develop strategies to document the intended program. Scheirer (1987) suggested that the investigation of the implementation process should be conducted separately from the testing of the program content itself, an analytic distinction that separates *what* the program does from the *way* the program was actually delivered. This does not preclude the importance of program theory to implementation; indeed, a theoretical basis is necessary to explicate the components of the program that could serve to compare the program as planned to the program as implemented. Scheirer (1987) further distinguished two aspects of implementation, involving the *extent* to which the program is delivered as planned and the *processes* that might explain implementation success or failure. Extent of implementation includes measuring both the accuracy of service provision and the scope of the program. Implementation process theory accounts for factors that affect the delivery of services, particularly within the interorganizational context of the program itself and the effects of the external environment.

Monitoring the fidelity of the treatment implementation is simpler when the roles of evaluator and treatment administrator are the same (Rezmovic, 1984). Henggeler and colleagues (Henggeler & Borduin, 1990, 1995; Henggeler, Melton, Smith, Schoenwald, & Hanley, 1993; Henggeler et al., 1995) participated in supervisory activities with the direct service providers in a study of multisystemic family preservation therapy for juvenile offenders and their families conducted in community settings. Through this process of intensive supervision and case review, the authors were able to assess the actual implementation of the treatment protocol by individual therapists. Additionally, the high level of direct involvement allowed the authors to make analytic decisions, omitting cases from therapists who consistently failed to follow the treatment protocol (Scherer, Brondino, Henggeler, Melton, and Hanley, 1994). However, there are significant drawbacks to the investigators performing summative evaluations of their own programs. Concerns about objectivity should require independent replication of findings. The task of investigating implementation issues is made more difficult in multidimensional service settings, such as in evaluations of overall systems of care for children and their families. Without the opportunity for the high level of direct involvement described above, evaluators must rely on other sources of information, including participants, service providers, program administrators, policy makers, and community organizations. Heflinger (1993, 1996a) described a case study approach using multiple methods and multiple sources of information used in the implementation study of the FBEP. This methodology included review of program documentation, such as committee meeting minutes and administrative reports; semistructured interviews with key informants, such as program staff, policy makers, and other agency staff; descriptive information about participants; and a network analysis (see Heflinger, 1996b), using both questionnaires and interviews with service providers and service agencies. These measures provided a comprehensive description of the project and demonstrated that a continuum of care was implemented with coordinated, family-focused, and individualized services made available for the children and families served (Heflinger, 1993, 1996a,b).

Friedman (1997) suggested that a more direct focus is needed to know whether changes implemented at the system level are translating into changes at the client level. In evaluations of both the Urban Child Mental Health Initiative (Friedman & Hernandez, 1993) and the Comprehensive Mental Health Services Program for Children (MACRO International and University of South Florida, 1995), structured interviews were conducted by trained professionals with the key individuals involved with a random sample of participants, including the children and their caregivers, case managers, and therapists. In addition, focus groups were held with multiple stakeholders in order to provide complementary information. Such methodologies appear promising for examining the interaction of the child and family with the overall service delivery system, although these procedures are still in the relatively early stages of development.

It is important to remember that the type of implementation study attempted will be informed by the type of implementation intended for a particular program. For example, Palumbo and Oliverio (1989) concluded that there are at least four theories of implementation, with differential implications for external validity. These include: (1) the top-down approach, where implementation occurs in a centralized organization with a high level of administrative control; (2) the bottom-up approach, where it is acknowledged that "street-level" providers have considerable discretion during implementation; (3) adaptive implementation and (4) evolutionary implementation, where programs must change while being implemented to adapt to local conditions or make progressive improvements during each new application. Thus, the challenge to researchers evaluating programs in the human service delivery field is to avoid a single method of assessing implementation, acknowledging that program goals and practice standards are disseminated along several levels, from administrators and policy makers to direct service providers to service recipients. Similarly, the developmental nature of mental health programs implemented in the real world indicates that evaluations should be ongoing, with the incorporation of feedback mechanisms between the research project and the program to report on barriers to care and solutions to such problems as they occur.

Recruitment and Eligibility

Another aspect of monitoring the fidelity of the evaluation is recruitment monitoring. The issue of slow recruitment of research participants has been discussed extensively in the literature (e.g., Lipsey, 1989, 1990; Sechrest, West, Phillips, Redner, & Yeaton, 1979), and its persistence as a problem, even with veteran service researchers, should serve as a caution to all evaluators. Slow recruitment can have serious implications for both internal validity and statistical power and highlights the critical importance of consistent monitoring by research staff. Thus, the process of determining eligibility of participants is an important component in determining the feasibility of implementing any program. Before beginning evaluation efforts, researchers should consider the consequences of insufficient recruitment, have plans in place to address potential problems, and have protocols established for the ongoing monitoring of recruitment.

When recruitment is slow, investigators can extend the time to obtain an adequate sample but at a high cost in usually limited resources, such as time and money. Additionally, the integrity of the study design itself can be threatened, particularly in designs using random assignment. Another approach, not one that is recommended, is to change the eligibility requirements for participation. A com-

mon strategy is to reduce requirements related to the severity of the disorder to increase the sample size. This strategy can alter the external validity of the study in significant ways. For example, as previously mentioned, the SCEP (Bickman et al., 1997c) had reduced external validity because of the exclusion of emergency cases, which was considered necessary for ethical reasons. However, since the study was intended to explore the effectiveness of a system of care for seriously emotionally disturbed (SED) children, external validity would have been further compromised by the inclusion of children below the threshold or criteria set for SED. By limiting the sample to nonemergency cases, the results may not be generalized to children with more severe problems. Similarly, if the SCEP had included (but did not) less severe cases (e.g., non-SED), the intervention of interest may have appeared more effective as a result of what is commonly known as "creaming," or selecting those children that may be more responsive to any treatment.

At the extreme, slow recruitment can lead to inadequate sample size. In evaluation research, attention to factors that affect statistical power is crucial, since type II errors (concluding that there was no effect when there actually was one) appear to be more common than type I errors (concluding there was an effect when there actually was not). A detailed discussion of this issue and the factors affecting the statistical power of evaluation designs, such as effect size, type of statistics used, sample size, and alpha level, is available elsewhere (see Bickman, 1990; Kraemer & Thiemann, 1987; Lipsey, 1989, 1990). However, the point should be made here that the consequences of inadequate evaluation design are critical, since insufficient statistical power can lead to null findings, making it appear that a program is ineffective and affecting the likelihood for funding and other future support (Bickman, 1990). Since the most visible factor that affects statistical power is the sample size, one of the first questions that evaluators need to address is the number of participants needed. Consequently, methods of ensuring sufficient recruitment while planning and during the period of evaluation are crucial.

A preliminary "pipeline study" can be useful in alleviating the problems associated with common delays in field experiments because the actual rate of recruitment of participants is usually substantially fewer than anticipated during the research design phase. Verifying client flow and availability with a pipeline study during the planning phase of a study would maintain statistical power by enabling evaluators to obtain nonoptimistic and accurate estimates of the pool of potential participants. Bickman et al. (1997c) set forth guidelines for conducting pipeline studies in their description of the baseline results of the SCEP. Information to be obtained from the program for a pipeline study includes: (1) the number of persons that apply or are referred or are recruited per month, (2) the relevant characteristics of those persons, (3) the eligibility criteria for the program, and (4) the number of new clients per month that the program can admit. This information would then be combined with eligibility criteria for the study and estimates of participation rates to provide an informed assessment about the number of eligible clients in the pipeline. However, key information is usually missing and must be estimated. In the SCEP, even though the researchers were aware of the problems, they underestimated recruitment problems because they did not have good data to judge how many children would be excluded from the study. They did not have an accurate way to estimate the number of children who would be excluded if they had siblings already receiving services, lived out of the county, and other such factors that had important influences on recruitment.

Because the SCEP conducted an evaluation of a well-established existing pro-

gram, capacity of the program was known and some but not all screening information was available from the records kept by the agencies involved. However, new programs may not have access to the necessary information needed for a pipeline study. The Assessing Coordinated Care Study (ACC) (Burns, Farmer, Angold, Costello, & Behar, 1996; Farmer et al., 1997; Morrissey, Johnsen, & Calloway, 1997) evaluated the impact of the North Carolina Children's Initiative, a new system of children's mental health care emphasizing coordination and case management. The problems of slow recruitment encountered in the ACC were illustrative of common issues, including loss of recruitment due to a delay between the start-up of the program and the beginning of the study, underestimation of the number of participants not eligible for the study, and slow intake of clients due to program characteristics. Farmer et al. (1997) suggested that evaluation studies should be designed to overestimate the number of participants needed by as much as two to three times the minimal estimate. This can be done by building in additional study sites or additional recruitment time during the design phase. Of course, until funding agencies are willing to support these increased costs and investigators are willing to admit to these problems, the field will continue to suffer from under-powered studies.

Aside from planning for insufficient recruitment by conducting a pipeline study or overestimating recruitment at the beginning of an evaluation, it is critical to monitor the process of recruitment after the study is in place. Building in protocols to monitor recruitment as part of the implementation monitoring provides the mechanism to identify and correct for unanticipated reasons for recruitment difficulties. For both studies mentioned above, the SCEP and the ACC, underestimation of ineligible clients was a large factor in the slow recruitment rates reported (Bickman et al., 1997c; Farmer et al., 1997). In addition, both were randomized studies, the nature of which requires particular attention to the assignment of participants to control and treatment groups. Without understanding and support, referring agencies' cooperation in the random assignment procedure may be limited. In both studies, slow recruitment rates were traced to miscommunication with agency staff or resistance to the randomization process. In field experiments, miscommunication often refers to the conflict between research protocol and normal clinical practice. Consequently, the importance of the research project itself should be explained and concessions may be made in order to respond to concerns related to clinical decision making. Although such problems cannot be avoided completely, they can be addressed by emphasizing a collaborative working environment between researchers and program staff and providing sufficient time to explore and create understanding of the concept behind the research design before the study begins. This would set the stage for such communication to continue during the evaluation process, leading to greater opportunity to address concerns as they occur.

Issues of Using Referred Populations

Random assignment of participants to treatment and control groups typically has been viewed as the "gold standard" in impact evaluation. It is the strongest design in terms of controlling threats to internal validity, yet it is difficult to implement in applied settings. Use of randomized designs requires adequate longi-

tudinal funding and sufficient start-up time to address procedural problems as well as to assure that the intervention is adequately developed and stable. As mentioned previously, the use of randomized designs typically raises concerns from stakeholders. Although there might be fewer objections and more willingness to be involved in randomized research from high-level administrators and policy makers (e.g., Bickman et al., 1997c), in the ACC Study, Farmer et al. (1997) found more intense feelings and distrust of randomization from clinical providers and staff, those dealing face-to-face with children with emotional and behavior disorders and their families. They speculated that such concerns most likely were due to the design's interference with normal clinical practice, the narrowing of clinical decision making, and the shift in authority and power. This section will highlight some of the barriers to randomization and practical guidelines for use in implementing randomized experiments (for a more thorough discussion regarding the use of randomized designs in applied research, see, e.g., Bickman, 1992; Hoagwood, 1994; Boruch, 1997; West & Sagaron, *in press*).

The most common objections to randomized experiments concern legal and ethical issues of denying treatment to children who are in need of services. While the use of random assignment has been ruled legal when challenged in the courts (Boruch & Wothke, 1985), evaluators need to be careful that none of the services a child might not receive as a part of the control group can be considered a service to which the child is legally entitled, such as a service funded under Medicaid. Addressing the ethical concerns of random assignment are more complicated, given that stakeholders in a given program are likely to hold strong opinions about the efficacy of their particular treatment. One counterargument to this position is that it may not be ethical to treat individuals with an intervention that has not yet demonstrated its effectiveness. However, ethical objections are especially valid if a child is denied treatment of any kind. Thus, it is advisable to compare a new treatment or program to already existing treatments rather than to a no-treatment control group. The disadvantage of this comparison is that there is likely to be a smaller effect associated with comparing two treatments than when comparing treatment versus no treatment. Thus, the statistical power of the study has to be sufficient to detect these smaller effects.

Respecting the role of stakeholders in decision making also presents ethical concerns. Because clinicians and program staff may object to loss of control to treat children as they see fit, particularly with children who are perceived as most in need, exceptions to randomization should be built into the evaluation study at the start. For example, in the ACC study, clinicians had the option to withhold a child from inclusion in the study if they felt that randomization to the control group would present an immediate threat to the child's well-being (Farmer et al., 1997). As previously mentioned, in the SCEP study, children who were judged to be emergency cases were excluded both for ethical and design considerations (Bickman et al., 1997c). Such planned exemptions should not threaten the integrity of the randomization, but they do affect the external validity. Regardless, any planned or unplanned exceptions to randomization should be clearly documented.

The issue that does affect internal validity is the migration of participants from one condition to another. For example, in the SCEP study (Bickman et al., 1997c), a condition required by the agency was that children in the control group who had several out-of-home placements should be considered eligible to receive the com-

prehensive system services. This was a fairly objective criterion and was never applied. However, since all studies are voluntary, there is nothing to stop the family from dropping out of the study and applying for the experimental treatment, assuming that it is available. A few families in the SCEP did this, but they agreed to continue to participate in the data collection. It is necessary not only to assign participants at entry to different conditions, but also to keep them in those conditions throughout the data collection period. For longitudinal studies this may mean years. This factor makes waiting list control groups inappropriate because it is usually not feasible to keep people on a waiting list for several years. However, again the importance of documentation should be emphasized. Information collected for such "cross-over" cases should include the reasons for changing treatment and if possible follow-up information on the families' perceptions comparing the different treatments.

Little attention has been paid to the perspectives of children and youth with emotional and behavior problems and their families who have participated in randomized research studies. It is common practice to follow institutional review board guidelines by obtaining informed consent from participants prior to random assignment and providing full disclosure about the nature of the evaluation (i.e., the assignment to a treatment or a control group at random). However, randomization may be viewed as taking away the participant's freedom of choice and active role in decision making, contributing to the view that the traditional research paradigm views children and their families as "objects" of study (Heflinger, 1987).

Given the problems of implementing randomized designs and the ethical concerns noted above, more attention should be devoted to alternative research methods. Friedman (1997) pointed out that randomized designs should be reserved for special situations and should not be used in initial attempts to evaluate services. Moreover, he contends that except in rare cases, randomized designs cannot be used to investigate the effectiveness of communitywide systems of care or initiatives, thus presenting special challenges to evaluators. It is important to note that multiple methodologies exist for conducting viable and useful evaluations without employing traditional scientific methodology such as randomization (e.g., Coulton, 1995; Weiss, 1995; Weiss & Greene, 1995; Yin, 1994). These include both quantitative methods, such as quasi-experiments, and qualitative methods, such as case studies.

Perceptions of Multiple Stakeholders

The active involvement of multiple stakeholders, those whose lives directly are affected by the program, services, or system, has been advocated as a mechanism to increase the validity and utility of evaluation results (e.g., Weiss, 1983a,b). Stakeholders should be defined broadly to include all those whose lives are affected by programs, services, and the evaluation of those services. Thus, those who have a stake in mental health services include (but are not limited to) the following: (1) policy makers, at the federal, state, and local levels; (2) practitioners, acknowledging differences among professionals in educational training and contact with clients; (3) administrators, including managed care professionals, upper management of care facilities, and purchasers; and (4) clients, again, including and acknowledging potential differences in perceptions among patients, their family

members, and members of the general public who are not currently receiving services.

The importance of cultivating collaborative relationships among evaluators, program staff, and service providers already has been discussed, particularly as it is relevant to preserving the integrity of the evaluation. Ensuring sufficient time and opportunity to meet prior to start-up and throughout the course of the evaluation can serve to reduce subsequent conflicts between researchers and program staff. Several concrete mechanisms are available for increasing collaboration, including maintaining an on-site research coordinator, conducting site visits, holding focus groups, and including multiple stakeholders as participants on program-related advisory committees (e.g., Bickman et al., 1997c; Evans & Armstrong, 1997; Farmer et al., 1997). Such mechanisms not only facilitate communication, making researchers more aware of clinical reality, but also give the research project "a face."

Fostering the involvement of multiple stakeholder groups in evaluation research goes beyond simply preserving internal validity. The development of the research questions themselves and assessment methods can be informed by a variety of perspectives. For example, the FBEP (Bickman et al., 1995a; Bickman, Summerfelt, & Bryant, 1995b) used a formalized brainstorming technique, called *concept mapping* (Trochim, 1989), to allow various stakeholder groups to define the key elements of quality in that mental health service delivery system. Key elements were given operational definitions and measured by various stakeholders as part of the assessment of the quality of services delivered within the program. Hernandez and colleagues (1998) have proposed an innovative information–feedback system that involves stakeholders in utilizing outcomes to improve delivery of services and increase their involvement in decision–making processes.

With the increasing push to control the quality of health services through market competition, more recent research has focused on assessing quality of care from the viewpoint of different stakeholders, particularly consumers of health services, rather than from the traditional clinician viewpoint. Following the lead of evaluation researchers and systems theorists (e.g., Ackoff, 1994), the best picture of quality may be gleaned from the input of multiple stakeholders, each experiencing different parts of the service system. Input from multiple stakeholders in the evaluation of quality theoretically increases the chance that the ultimate goal of improving the quality of care will be realized by increasing the validity, fairness, range, and responsiveness on which program decisions are based (Weiss, 1983a,b).

However, research typically has not included all relevant stakeholders in the study of quality of care, and thus has not incorporated a true systems framework, with the vast majority of efforts examining only one or two perspectives simultaneously. Moreover, most research has addressed quality of health and mental health care indirectly by exploring stakeholders' views of physician or clinic performance (e.g., Coser, 1956; Hobbs, Perrin, & Ireys, 1985; Lohr, Donaldson, & Walker, 1991) or by investigating client satisfaction (e.g., Hall & Dornan, 1988; Kurata, Nogawa, Phillips, Hoffman, & Werblun, 1992; Ware & Davies, 1983). Yet, recent research has explored clients' desire for, understanding of, and use of disseminated information on quality (e.g., Agency for Health Care Policy, 1995; Hibbard & Jewett, 1995; National Committee on Quality Assurance, 1995). In 1994, a multistakeholder conference including clients, families, clinicians, managed care representatives, and insurers among others was convened by the National Alliance

for the Mentally Ill, Johns Hopkins University, and the NIMH. More theoretically based work is beginning to appear in the literature that promotes a "multiconstituency" view (Connolly, Conlon, & Deutsch, 1980) of service quality (e.g., Brown & Swartz, 1989; Parasuraman, Zeithaml, & Berry, 1985; Wilde, Starrin, Larsson, & Larsson, 1993); yet clearly, much work remains to be done.

Chelimsky (1995) emphasized the importance of involving children with emotional and behavior disorders and their families as coparticipants rather than as "targets" in research studies. They, as well as other relevant stakeholders, are critical in helping to define the problems that plague them and in designing, implementing, and evaluating interventions. There are many options available for collaboration, from including the views of national and community family-oriented groups in the formulation of general intervention goals to using focus groups to obtain the views of children and families in the formulation of specific evaluation studies. Further, the participation of program staff and providers promotes the evaluators' understanding of the "users' milieu" (Chelimsky, 1995, p. 8), thereby increasing the relevance, breadth, and utility of evaluation results. Bringing a multitude of stakeholders actively into the research process, thus providing a broad base of support, will increase the likelihood that evaluation results will be used in decisionmaking.

INTERPRETING NULL FINDINGS AND THEIR IMPLICATIONS FOR SERVICES RESEARCH

The likelihood of finding no effects in mental health services research (e.g., Bickman, 1990, 1992) demands stringent attention to methodological factors that might lead to the failure to detect an actual difference, as discussed in the previous sections. When no difference is found in well-designed and well-implemented evaluation studies, the implications for the field can be tremendous. Sechrest and Walsh (1997) stated that when public policy is in the balance, in order to avoid ambiguity, such conclusions should be stated as "acceptance of the null hypothesis" rather than the usual "failure to reject" (p. 537).

The recent publication of the results of the FBEP (Bickman et al., 1995a; Bickman, 1995a,b, 1997) has engendered a widespread debate about the effectiveness of a continuum of care for children with emotional and behavior problems and their families.* Seen as a methodologically sound study even by its critics (Behar, 1997; Friedman & Burns, 1996), the FBEP tested the theory "that a comprehensive, integrated, and coordinated continuum of care is more cost-effective than a fragmented service system with a limited variety of services" (Bickman, 1996a, p. 690). Although it was found that children in the treatment group had greater access to services, children at both the treatment site and the comparison sites showed similar levels of improvement in outcomes. This "acceptance of the null hypothesis," along with the finding that costs were significantly higher at the treatment site, has called into question strongly held beliefs about the nature of children's mental health services.

*For more information on the FBEP and responses to the project, see a special section "Psychology in the Public Forum," in *American Psychologist* (1997, Vol. 52(5), pp. 536-565); and a special issue of the *Journal of Mental Health Administration* (1996, Vol. 23).

In a commentary about the findings of the FBEP, Weisz, Han, and Valeri (1997) pointed out that the mental health field does not have the technology to adequately describe treatment, which limits the ability to describe and improve the quality of services in the real world. Bickman (1997) suggested two complementary approaches to enhance quality of care at both the treatment level and the systems level. The *diagnostic-specific treatment improvement approach* calls for the development of manualized or standardized treatments for specific diagnostic conditions (Hoagwood et al., 1995; Chambless & Hollon, 1998) that once validated in laboratory settings are then tested in typical community settings. The *generic services improvement approach* emphasizes improving general elements of services (e.g., comprehensive assessment, supervision, treatment planning, concurrent assessment) that are believed to constitute quality treatment for any disorder.

This corresponds to the argument that the field is in dire need of more basic research, such as randomized trials (Arnold, Hoagwood, Jensen, & Vitiello, 1997; Burns & Friedman, 1990), in addition to the evaluation of continuums and systems of care. Chelimsky (1995) has reasoned that the impact of such macro-level studies on public policy will be diminished until the fundamental link between an intervention and its putative effect has been established. On the other hand, some evaluators (e.g., Cross & Saxe, 1997) have referred to "micro-level" efforts as reductionistic and ultimately self-defeating, emphasizing the need to test the impact of the entire system. This is consistent with a systems perspective (e.g., Ackoff, 1994), in which the whole is more than the sum of its parts, highlighting the need to recognize that a service system has a synergy of its own.

However, rather than viewing the situation as a "catch 22," micro-level and macro-level research could be viewed as complementary approaches. While the limitations of each type of research should be acknowledged, such a multiple-methods approach may prove more informative and useful to improving children's mental health services. This coincides with the point made earlier in the chapter regarding the importance of combining efficacy and effectiveness research (Hoagwood et al., 1995; Weisz et al., 1995), with smaller, specific intervention trials leading to larger demonstration studies. Such a model allows for a more in-depth investigation of the efficacy of an intervention and for the evaluation of the transfer of such interventions into community settings. For example, Henggeler and colleagues (1995) stated that the successful transfer of multisystemic therapy for juvenile delinquents from the 'laboratory' to the 'clinic' was due not only to a strong theoretical and empirical foundation but also to careful attention to service delivery.

ENHANCING THE CLINICAL RELEVANCE OF MENTAL HEALTH SERVICES RESEARCH

As stated previously, the ultimate purpose of children's mental health services research is to determine whether or not a program or system of care contributed to positive outcomes for those children and families being served. Findings from well-designed evaluation studies are intended to inform the future direction of the field, influencing both policy decisions for funding and the improvement of mental health services and clinical decision making in the delivery of mental health

services. However, evaluation has generally been criticized as narrow and irrelevant, unresponsive to the information needs of program staff, unfair, not considering the needs of service recipients, and unused, often not influencing actual decision making about services (Weiss, 1983a). Increasing the usefulness of evaluation research requires more attention to issues that would enhance clinical relevance, including both increased involvement of clinicians in the determination of research issues and greater understanding of the impact of treatment research on clinician behavior.

The importance of collaborative relationships between evaluators and multiple groups of stakeholders already has been discussed in terms of preserving the integrity of the research design (e.g., Bickman et al., 1997c; Farmer et al., 1997) and developing specific assessment measures (e.g., Bickman et al., 1995a,b). However, further discussion is warranted about the value of including the perspectives of clinicians and other direct service providers in determining what are the relevant questions. Enhancing the investment of nonresearch staff in the project means promoting active participation, being cautious that the facilitation of communication does not become a "one-way street" for the dissemination of program information from researchers and administrators to direct service providers. Because of the compromises in design that could stem from such collaboration, the onus is on the researchers to devise flexible methodologies and research strategies, while also determining at the outset what components of the research design are essential, and thus not subject to change. Such an attitudinal change would reflect the understanding that mental health services research in the real world is dynamic, and may entail a trade-off of traditional methodological considerations for more relevant, collaborative research. Instead of viewing this as a barrier to conducting research in applied settings, it could be conceptualized as a challenge to expand evaluation methodology.

Similarly, an attitudinal change on the part of clinicians also is necessary in order to further the field of children's mental health services research. Strong research findings do not guarantee that research-based treatment will be used in applied clinical settings. It has been suggested that "the gulf that divides clinical practice and clinical research is now accepted as a fact of life by many in the mental health professions and in academia" (Weisz et al., 1995, p. 688). Indeed, it has been found that clinicians usually do not view research as relevant to their daily clinical practice (Cohen, Sargent, & Sechrest, 1986; Morrow-Bradley & Elliott, 1986; Raw, 1993). Weisz et al. (1995) identified and addressed some of the obstacles to linking research and clinical practice, including the beliefs that psychotherapy is not a scientific endeavor, that outcome research is irrelevant to the typical complexity of cases seen in clinics, and a devotion to personally appealing but untested theoretical orientations. More practical concerns include the possible lack of funding for restructuring clinic procedures and the difficulty accessing and using proven interventions for the diverse population of children typically seen in clinics.

Many of these concerns are well-justified, stimulating the need for more researchers to address issues related to the "transportability" of research-based treatment to community settings (Kazdin & Kendall, 1998; Weiss, Catron, Harris, & Phung, 1999), as demonstrated successfully with Henggeler and colleagues research on multisystemic therapy (Henggeler et al., 1995). Moreover, increasing the accessibility of research findings needs to be emphasized and promoted through a

variety of means, such as manual-based treatment, professional treatment guidelines, and comprehensive resource books (e.g., Hibbs & Jenson, 1996; Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998; National Institute of Mental Health, 1999). However, it also is the responsibility of clinicians to avail themselves of such knowledge and to take an active voice in directing future research questions. The field of children's mental health services can only benefit from a closer collaboration between clinicians and researchers. As Weisz et al. (1995) concluded, "if the obstacles to research-clinician collaboration can be overcome, both groups may profit, and to the ultimate benefit of the children and families who seek help" (p. 699).

IMPLICATIONS AND FUTURE DIRECTIONS

Estimates suggest that 12–22% of children and youth suffer from a diagnosable mental disorder (Brandenburg, Friedman, & Silver, 1990; Costello, 1989), with a possible increase in prevalence over the last few years (Stroul, 1996). Yet, it appears that only a small proportion of those in need of services receive them, particularly in nonmental health settings such as schools, primary health care, and juvenile justice facilities (Burns, 1990; Hoagwood, 1994; Saxe, Cross, Silverman, Batchelor, & Dougherty, 1987). Well-designed, well-implemented, and well-analyzed research is needed to inform public policy and further theoretical advances in the field of children's mental health services. Friedman (1997) cautioned that without such controlled studies, there might be a decrease in funding for existing services as competition for limited resources increases. As it is, funding is often inadequate for efforts such as research synthesis (e.g., meta-analysis), even though they have proved to be a critical tool in assessing the effectiveness of psychological interventions (Lipsey & Wilson, 1993; Windle, 1994).

In order to improve children's mental health services, more research on specific interventions is needed in addition to evaluations of service systems in applied settings. Micro-level studies can be useful to investigating the effectiveness of clinical interventions, broadening the picture of the specific components within the "black box" that are related to outcome. Because the establishment of intervention efficacy is no guarantee of effectiveness in community settings, macro-level systems research can primarily address the implementation of programs and systems-level reforms in natural settings.

To accomplish these goals, graduate programs in a variety of disciplines (e.g., psychology, psychiatry, economic, social work, nursing, public policy) should include training in mental health service research methods. Courses in program evaluation, quasi-experimental design, and meta-analysis as well as exposure to the literature regarding micro- and macro-level research would further understanding of the issues involved in evaluation research. In addition, there must be more active involvement of multiple groups of stakeholders, including policy makers, administrative officials, researchers, practitioners, and clients. This may serve to increase the relevance and support for evaluation studies and the use of results in decision making. Moreover, since the knowledge that is needed in this field is so vast and the research resources are so small, there needs to be a major shift from service organizations devoting all their resources to service delivery to the inclu-

sion of services research. These organizations need to become learning organizations in that they learn what works in their own setting. In addition, because of the increasing emphasis on managed care, this industry should take more responsibility for supporting research and development (Bickman, 1996a). Such a focus need not be discordant with current management strategies (e.g., cutting costs), but may be viewed as complementary to the common goal of understanding how clinical practice patterns (i.e., managing care) affect children's mental health outcomes (e.g., Ogles, Trout, Gillespie, & Penkert, 1998; Pires, Stroul, & Armstrong, 1998; Stroul, Pires, Armstrong, & Myers, 1998). This orientation would mesh well with the increasing emphasis on evaluation in the public sector, such as the requirement that states with health care demonstrations undertake external evaluations as per the Health Care Financing Administration, the federal agency that administers Medicare and Medicaid programs. With the ultimate goal of improving the lives of children and families in need, researchers must take on the responsibility of innovation and creativity in evaluation and open channels of communication with stakeholders in order to further the field of children's mental health services research.

ACKNOWLEDGMENTS. The writing of this chapter was supported in part by grant RO1-MH-46135 and training grant T32MH-19544 from the National Institute of Mental Health to Dr. Leonard Bickman. Many thanks are due to Sandra Jo Wilson for her helpful comments on an earlier draft of this chapter.

REFERENCES

- Ackoff, R. L. (1994). *The democratic corporation: A radical prescription for recreating corporate America and rediscovering success*. New York: Oxford University Press.
- Agency for Health Care Policy and Research. (1995). *Conference summary: Consumer survey information in a reforming health care system* (AHCPR Publication No. 95-0083). Rockville, MD: US Department of Health and Human Services.
- Arnold, L. E., Hoagwood, K., Jensen, P. S., & Vitiello, B. (1997). Towards clinically relevant trials. *Psychopharmacology Bulletin*, 33(1), 135-142.
- Behar, L. B. (1997). The Fort Bragg Evaluation: A snapshot in time. *American Psychologist*, 52(5), 557-559.
- Bickman, L. (Ed.). (1987). *Using program theory in evaluation: Vol. 33. New directions for program evaluation*. San Francisco: Jossey-Bass.
- Bickman, L. (1990). Study design. In Y. T. Yuan & M. Rivest (Eds.), *Preserving families: Evaluation resources for practitioners and policymakers* (pp. 132-166). Newbury Park, CA: Sage.
- Bickman, L. (1992). Designing outcome evaluations for children's mental health services: Improving internal validity. In L. Bickman & D. J. Rog (Eds.), *Evaluating mental health services for children. Vol. 54. New directions for program evaluation* (pp. 57-68). San Francisco: Jossey-Bass.
- Bickman, L. (1995, March). *The Fort Bragg Managed Care Experiment: Findings and implications*. Paper presented at the Eighth Annual Research Conference, A System of Care for Children's Mental Health. Expanding the Research Base, Florida Mental Health Institute, Tampa, FL.
- Bickman, L. (1996a). A continuum of care: More is not always better. *American Psychologist*, 51, 689-701.
- Bickman, L. (1996b). The application of program theory to the evaluation of a managed mental health care system. *Evaluation and Program Planning*, 19(2), 111-119.
- Bickman, L. (Ed.). (1996c). The Fort Bragg Experiment [Special Issue]. *Journal of Mental Health Administration*, 23(1).
- Bickman, L. (1997). Resolving issues raised by the Fort Bragg evaluation: New directions for mental health services research. *American Psychologist*, 52(5), 562-565.

- Bickman, L., Guthrie, P. R., Foster, E. M., Lambert, E. W., Summerfelt, W. T., Breda, C. S., & Heflinger, C. A. (1995a). *Evaluating managed mental health services: The Fort Bragg experiment*. New York: Plenum Press.
- Bickman, L., Lambert, W., Northrup, D., Salzer, M., & Summerfelt, W. T. (1997a, May). *A conceptually-based adolescent outcome monitoring system with empirical foundations* (Report to Compass Information Systems, Inc.). Nashville, TN: Vanderbilt University.
- Bickman, L., & Rog, D. (Eds.). (1997). *Handbook of applied social research methods*. Thousand Oaks, CA: Sage.
- Bickman, L., Rog, D., & Hedrick, T. E. (1997b). Applied research design: A practical approach. In L. Bickman & D. Rog (Eds.), *Handbook of applied social research methods* (pp. 5–38). Newbury Park, CA: Sage.
- Bickman, L., & Salzer, M. S. (1997). Measuring quality in mental health services. *Evaluation Review*, 21(3), 285–291.
- Bickman, L., Summerfelt, W. T., & Bryant, D. (1995b). The quality of services at the Fort Bragg Demonstration. *The Journal of Mental Health Administration*, 23(1), 30–39.
- Bickman, L., Summerfelt, W. T., Firth, J. M., & Douglas, S. M. (1997c). The Stark County Evaluation Project: Baseline results of a randomized experiment. In L. Bickman (Series Ed.), C. Nixon, & D. A. Northrup (Vol. Eds.), *Children's mental health services: Vol. 3. Evaluating mental health services: How do programs for children "work" in the real world?* (pp. 209–230). Thousand Oaks, CA: Sage.
- Bickman, L., Summerfelt, W. T., & Foster, M. (1996). Research on systems of care: Implications of the Fort Bragg Evaluation. In B. A. Stroul (Ed.), *Children's mental health: Creating systems of care in a changing society* (pp. 337–355). Baltimore: Paul H. Brookes.
- Bickman, L., Summerfelt, W. T., & Noser, K. (1997d). Comparative outcomes of emotionally disturbed children and adolescents in a system of services and usual care. *Psychiatric Services*, 48(12), 1543–1548.
- Boruch, R. F. (1997). *Randomized experiments for planning and evaluation: A practical guide*. In L. Bickman & D. J. Rog (Series Eds.), *Applied social research methods* (Vol. 44). Thousand Oaks, CA: Sage.
- Boruch, R. F., & Gomez, H. (1977). Sensitivity, bias, and theory in impact evaluation. *Professional Psychology*, 8, 411–434.
- Boruch, R. F., & Wotheke, W. (Eds.). (1985). *Randomization and field experimentation: Vol. 28. New directions for program evaluation*. San Francisco: Jossey-Bass.
- Brandenburg, N., Friedman, R., & Silver, S. (1990). The epidemiology of childhood psychiatric disorders: Prevalence findings from recent studies. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 76–83.
- Brown, S. W., & Swartz, T. A. (1989). A gap analysis of professional service quality. *Journal of Marketing*, 53, 92–98.
- Burns, B. J. (1990). Mental health service use by adolescents in the 1970s and 1980s. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 144–149.
- Burns, B. J. (1994). The challenges of child mental health services research. *Journal of Emotional and Behavioral Disorders*, 2, 254–259.
- Burns, B. J., Farmer, E. M. Z., Angold, A., Costello, E. J., & Behar, L. B. (1996). A randomized trial of case management for youths with serious emotional disturbance. *Journal of Clinical Child Psychology*, 25(4), 476–486.
- Burns, B. J., & Friedman, R. M. (1990). Examining the research base for child mental health services and policy. *Journal of Mental Health Administration*, 17, 87–98.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66(1), 7–18.
- Chelimsky, E. (1995). Preamble: New directions in evaluation. In R. Picciotto & R. C. Rist (Eds.), *Evaluating country development policies and programs: New approaches for a new agenda*: Vol. 67. *New directions for evaluations* (pp. 3–12). San Francisco: Jossey-Bass.
- Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chen, H., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review*, 7(3), 283–302.
- Cohen, L., Sargent, M., & Sechrest, L. (1986). Use of psychotherapy research by professional psychologists. *American Psychologist*, 41, 198–206.

- Connolly, T., Conlon, E. J., & Deutsch, S. J. (1980). Organisational effectiveness: A multiple-constituency approach. *Academy of Management Review*, 5, 211–217.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Coser, R. L. (1956). A home away from home. *Social Problems*, 4, 3–17.
- Costello, E. J. (1989). Developments in child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 28, 836–841.
- Coulton, C. J. (1995). Using community-level indicators of children's well-being in comprehensive community initiatives. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 173–199). Washington, DC: The Aspen Institute.
- Cronbach, L. J., Ambro, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F., & Weiner, S. S. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Cross, T. P., & Saxe, L. (1997). Many hands make mental health systems of care a reality: Lessons from the mental health services program for youth. In L. Bickman (Series Ed.), C. Nixon, & D. A. Northrup (Vol. Eds.), *Children's mental health services: Vol. 3. Evaluating mental health services: How do programs for children "work" in the real world?* (pp. 45–72). Thousand Oaks, CA: Sage.
- Evans, M. E., & Armstrong, M. I. (1997). The development of a state's perspective on keeping children in the community: Strategies for change. In L. Bickman (Series Ed.), C. Nixon, & D. A. Northrup (Vol. Eds.), *Children's mental health services: Vol. 3. Evaluating mental health services: How do programs for children "work" in the real world?* (pp. 73–94). Thousand Oaks, CA: Sage.
- Farmer, E. M. Z., Burns, B. J., Guiles, H. B., Behar, L., & Gerber, D. I. (1997). Conducting randomized clinical trials in children's mental health. In L. Bickman (Series Ed.), C. Nixon, & D. A. Northrup (Vol. Eds.), *Children's mental health services: Vol. 3. Evaluating mental health services: How do programs for children "work" in the real world?* (pp. 209–230). Thousand Oaks, CA: Sage.
- Friedman, R. M. (1997). Services and service delivery systems for children with serious emotional disorders: Issues in assessing effectiveness. In L. Bickman (Series Ed.), C. Nixon, & D. A. Northrup (Vol. Eds.), *Children's mental health services: Vol. 3. Evaluating mental health services: How do programs for children "work" in the real world?* (pp. 209–230). Thousand Oaks, CA: Sage.
- Friedman, R. M., & Burns, B. J. (1996). The evaluation of the Fort Bragg Demonstration Project: An alternative interpretation of the findings. *Journal of Mental Health Administration*, 21(1), 128–136.
- Friedman, R. M., & Hernandez, M. (1993, October). *Special challenges in evaluating multisite system reform efforts: A focus on mental health services*. Paper presented at the Fifteenth Annual Research Conference of the Association for Public Policy Analysis and Management, Washington, DC.
- Hall, J. A., & Dornan, M. C. (1988). What parents like about their medical care and how often they are asked: A meta-analysis of the satisfaction literature. *Social Science and Medicine*, 27, 935–939.
- Heflinger, C. A. (1987). Psychotherapy research ethics: Continuing the debate on controlled clinical trials. *American Psychologist*, 42(10), 956–957.
- Heflinger, C. A. (1993, October). *Final report of the implementation study of the Fort Bragg evaluation project*. Nashville, TN: Center for Mental Health Policy, Vanderbilt University.
- Heflinger, C. A. (1996a). Implementing a system of care: Findings from the Fort Bragg evaluation project. *Journal of Mental Health Administration*, 23(1), 16–29.
- Heflinger, C. A. (1996b). Measuring service system coordination in managed mental health care for children and youth. *Evaluation and Program Planning*, 19(2), 155–163.
- Henggeler, S. W., & Borduin, C. M. (1990). *Family therapy and beyond: A multisystemic approach to treating the behavior problems of children and adolescents*. Pacific Grove, CA: Brooks/Cole.
- Henggeler, S. W., & Borduin, C. M. (1995). Multisystemic treatment of serious juvenile offenders and their families. In I. M. Schwartz & P. AuClair (Eds.), *Home-based services for troubled children* (pp. 113–130). Lincoln: University of Nebraska Press.
- Henggeler, S. W., Melton, G. B., Smith, L. A., Schoenwald, S. K., & Hanley, J. H. (1993). Family preservation using multisystemic treatment: Long-term follow-up to a clinical trial with serious juvenile offenders. *Journal of Child and Family Studies*, 2, 283–293.
- Henggeler, S. W., Schoenwald, S. K., & Pickrel, S. G. (1995). Multisystemic therapy: Bridging the gap between university- and community-based treatment. *Journal of Consulting and Clinical Psychology*, 63, 709–717.
- Henggeler, S. W., Schoenwald, S. K., Pickrel, S. G., Rowland, M. D., & Santos, A. B. (1994). The contribution of treatment outcome research to the reform of children's mental health services: Multisystemic therapy as an example. *Journal of Mental Health Administration*, 21(3), 229–239.

- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic treatment of antisocial behavior in children and adolescents*. New York: Guilford Press.
- Hernandez, M., Hodges, S., & Cascardi, M. (1998). The ecology of outcomes: System accountability in children's mental health. *Journal of Behavioral Health Services and Research*, 25(2), 136–150.
- Hibbard, J. H., & Jewett, J. J. (1995, June). Using report cards to inform and empower: Consumer understanding of quality-of-care information. Paper presented at the Annual Meeting of the Association for Health Services Research.
- Hibbs, E. D., & Jensen, P. S. (Eds.) (1996). *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice*. Washington, DC: American Psychological Association.
- Hoagwood, K. (1994). Introduction to the special section: Issues in designing and implementing studies in non-mental health care sectors. *Journal of Clinical Child Psychology*, 23, 114–120.
- Hoagwood, K., Hibbs, E., Brent, D., & Jensen, P. (1995). Introduction to the special edition: Efficacy and effectiveness studies of child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology*, 63, 683–687.
- Hobbs, N., Perrin, J., & Ireys, H. (1985). *Chronically ill children and their families*. San Francisco: Jossey-Bass.
- Institute of Medicine. (1989). *Research on children and adolescents with mental, behavioral, and developmental disorders*. Washington, DC: National Academy Press.
- Kazdin, A. E., Bass, D., Ayers, W. A., & Rodgers, A. (1990). Empirical and clinical focus of child and adolescent psychotherapy research. *Journal of Consulting and Clinical Psychology*, 58, 729–740.
- Kazdin, A. E., & Kendall, P. C. (1998). Current progress and future plans for developing effective treatments: Comments and perspectives. *Journal of Clinical Child Psychology*, 27(2), 217–226.
- Kelleher, K., & Long, N. (1994). Barriers and new directions in mental health services research in the primary care setting. *Journal of Clinical Child Psychology*, 23, 133–142.
- Kendall, P. C., & Southam-Gerow, M. A. (1995). Issues in the transportability of treatment: The case of anxiety disorders in youth. *Journal of Consulting and Clinical Psychology*, 63, 702–708.
- Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*. Newbury Park, CA: Sage.
- Kurata, J. H., Nogawa, A. K., Phillips, D. M., Hoffman, S., & Werblun, M. N. (1992). Patient and provider satisfaction with medical care. *Journal of Family Practice*, 35, 176–179.
- Lambert, E. W., Nixon, C. T., Simpkins, G., & Bickman, L. (1996, December). Mental health outcomes for treated children: Definitions of improvement, effects of severity, and predicting clinical results. Final Report to the US Army MEDCOM contract DA-DA10-96-C-030. Nashville, TN: Vanderbilt University.
- Lipsey, M. W. (1989). *Design sensitivity: Statistical power for experimental research*. Newbury Park, CA: Sage.
- Lipsey, M. W. (1990). *Design sensitivity*. Newbury Park: Sage.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48(12), 1181–1209.
- Lohr, K. N., Donaldson, M., & Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. *Quality Review Bulletin*, 17, 242–253.
- Lonigan, C. J., Elbert, J. C., & Johnson, S. B. (1998). Empirically supported psychosocial interventions for children: An overview. *Journal of Clinical Child Psychology*, 27(2), 138–145.
- MACRO International and University of South Florida (1995). *Evaluation of the comprehensive mental health services program for children with severe emotional disturbances: Overview of the evaluation*. Atlanta, GA: MACRO International.
- Mechanic, D. (1996). Emerging issues in international mental health services research. *Psychiatric Services*, 47(4), 371–375.
- Mechanic, D., Bevilacqua, J. J., Goldman, H., Hargreaves, W., Howe, J., Knisley, M., Scherl, D. J., Stuart, G., Unhjem, M. B., & Lalley, T. L. (1992). Research resources. *Schizophrenia Bulletin*, 15, 669–696.
- Morrissey, J. P., Calloway, M., Bartko, W. T., Ridgely, M. S., Goldman, H., & Paulson, R. I. (1994). Local mental health authorities and service system change: Evidence from the Robert Wood Johnson Foundation Program on Chronic Mental Illness. *Milbank Quarterly*, 72, 1–32.
- Morrissey, J. P., Johnsen, M. C., & Calloway, M. O. (1997). Evaluating performance and change in mental health systems serving children and youth: An interorganizational network approach. *Journal of Mental Health Administration*, 24(1), 4–22.

- Morrow-Bradley, C., & Elliott, R. (1986). Utilization of psychotherapy research by practicing psychologists. *American Psychologist*, 41, 188–197.
- National Committee on Quality Assurance. (1995). *NCQA consumer information project focus group report: Executive summary*. Washington, DC: Author.
- National Institute of Mental Health (1990). *National plan for research on child and adolescent mental disorders*. Rockville, MD: National Institutes of Health.
- National Institute of Mental Health (1999). *Child and adolescent services research strategic planning meeting* (Belmont Conference Center Draft Report). Rockland, MA: National Institute of Mental Health, Division of Services and Intervention Research.
- Nixon, C. T. (1997). How does evaluation of mental health services for children work in the real world? In L. Bickman (Series Ed.), C. Nixon, & D. A. Northrup (Vol. Eds.), *Children's mental health services: Vol. 3. Evaluating mental health services: How do programs for children "work" in the real world?* (pp. 1–15). Thousand Oaks, CA: Sage.
- Ogles, B. M., Trout, S. C., Gillespie, D. K., & Penkert, K. S. (1998). Managed care as a platform for cross-system integration. *Journal of Behavioral Health Services and Research*, 25(3), 252–268.
- Palumbo, D. J., & Oliverio, A. (1989). Implementation theory and the theory-driven approach to validity. *Evaluation and Program Planning*, 12, 337–344.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49, 41–50.
- Persons, J. B. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy: A proposed remedy. *American Psychologist*, 46(2), 99–106.
- Pires, S. A., Stroul, B. A., & Armstrong, M. I. (1998). Impact of managed care in the public sector: Children and adolescents with behavioral health problems and their families. *Behavioral Healthcare Tomorrow*, 7(3), 41–46.
- Raw, S. D. (1993, March). Does psychotherapy research teach us anything about psychotherapy? *Behavior Therapist*, 75–76.
- Rezmovic, E. L. (1984). Assessing treatment implementation amid the slings and arrows of reality. *Evaluation Review*, 8, 187–204.
- Santos, A. B., Henggeler, S. W., Burns, B. J., Arana, G. W., & Meisler, N. (1995). Research on field-based services: Models for reform in the delivery of mental health care to populations with complex clinical problems. *American Journal of Psychiatry*, 152, 1111–1123.
- Saxe, L., Cross, T. P., Silverman, N., Batchelor, W. F., & Dougherty, D. (1987). *Children's mental health: Problems and treatment*. Durham, NC: Duke University Press.
- Scheirer, M. A. (1987). Program theory and implementation theory: Implications for evaluators. In L. Bickman (Ed.), *Using program theory in evaluation: Vol. 33. New directions for program evaluation* (pp. 19–42). San Francisco: Jossey-Bass.
- Scherer, D. G., Brondino, M. J., Henggeler, S. W., Melton, G. B., & Hanley, J. H. (1994). Multisystemic family preservation therapy: Preliminary findings from a study of rural and minority serious adolescent offenders. *Journal of Emotional and Behavioral Disorders*, 2, 198–206.
- Sechrest, L., & Walsh, M. (1997). Dogma or data: Bragging rights. *American Psychologist*, 52(5), 536–540.
- Sechrest, L. B., West, S. G., Phillips, M. A., Redner, R., & Yeaton, W. (1979). Some neglected problems in evaluation research: Strength and integrity of treatments. In L. Sechrest, S. G. West, M. A. Phillips, R. Redner, & W. Yeaton (Eds.), *Evaluation studies review annual* (Vol. 4, pp. 15–35). Beverly Hills, CA: Sage.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752–760.
- Speer, D. C. (1994). Can treatment research inform decision makers? Nonexperimental methods issues and examples among older outpatients. *Journal of Consulting and Clinical Psychology*, 62, 560–568.
- Speer, D. C., & Newman, F. L. (1996). Mental health services outcome evaluation. *Clinical Psychology: Science and Practice*, 3(2), 105–129.
- Starfield, B. (1973). Health services research: A working model. *New England Journal of Medicine*, 289(3), 132–136.
- Stroul, B. A. (1996). Introduction: Progress in children's mental health. In B. A. Stroul (Ed.), *Children's mental health: Creating systems of care in a changing society* (pp. xxi–xxxiii). Baltimore: Paul H. Brookes.
- Stroul, B. A., & Friedman, R. M. (1986). *A system of care for severely emotionally disturbed children and youth*. Washington, DC: Georgetown University Child Development Center, CASSP Technical Assistance Center.

- Stroul, B. A., & Friedman, R. M. (1996). The system of care concept and philosophy. In B. A. Stroul (Ed.), *Children's mental health: Creating systems of care in a changing society* (pp. 3–21). Baltimore: Paul H. Brookes.
- Stroul, B. A., Pires, S. A., Armstrong, M. I., & Meyers, J. C. (1998). The impact of managed care on mental health services for children and their families. *Future of Children*, 8(2), 119–133.
- Summerfelt, W. T. (1992). *The development of a computerized version of the Child Assessment Schedule: An initial study*. Unpublished master's thesis.
- Summerfelt, W. T., & Hodges, V. K. (1992, March). *Test-retest reliability of an interactive computerized interview with adolescents: The computerized Child Assessment Schedule*. Poster presented at the 5th Annual Research Conference, A System of Care for Children's Mental Health, Research and Training Center for Children's Mental Health, University of South Florida, Tampa, FL.
- Trochim, W. M. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, 12, 1–16.
- Ware, J. E., & Davies, A. R. (1983). Behavioral consequences of consumer dissatisfaction with medical care. *Evaluation and Program Planning*, 6, 291–297.
- Weiss, B., Catron, T., Harris, V., & Phung, T. M. (1999). The effectiveness of traditional child psychotherapy. *Journal of Consulting and Clinical Psychology*, 67(1), 82–94.
- Weiss, C. H. (1983a). The stakeholder approach to evaluation: Origins and promise. In A. S. Byrk (Ed.), *Stakeholder-based evaluation: Vol. 17. New directions for program evaluation* (pp. 3–14). San Francisco: Jossey-Bass.
- Weiss, C. H. (1983b). Toward the future of stakeholder approaches in evaluation. In A. S. Byrk (Ed.), *Stakeholder-based evaluation: Vol. 17. New directions for program evaluation* (pp. 83–96). San Francisco: Jossey-Bass.
- Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 65–92). Washington, DC: The Aspen Institute.
- Weiss, H. B., & Greene, J. C. (1995). An empowerment partnership for family support and education programs and evaluations. *Family Science Review*, 5, 131–148.
- Weisz, J. R., & Jensen, P. S. (1999). Efficacy and effectiveness of psychotherapy and psychopharmacotherapy with children and adolescent. *Mental Health Services Research*, 1(3), 125–157.
- 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, 688–701.
- Weisz, J. R., Han, S. S., & Valeri, S. M. (1997). More of what? Issues raised by the Fort Bragg study. *American Psychologist*, 52(5), 541–545.
- Weisz, J. R., Weiss, B., Aliche, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology*, 55, 542–549.
- Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinics: Effects of child and adolescent psychotherapy. *American Psychologist*, 47, 1578–1585.
- West, S. G., & Sagarin, B. J. (in press). Participant selection and loss in randomized experiments. In L. Bickman (Ed.), *Volume II: Contributions to research design: Donald Campbell's legacy*. Newbury Park: Sage.
- Wilde, B., Starrin, B., Larsson, G., & Larsson, M. (1993). Quality of care from a patient perspective: A grounded theory study. *Scandinavian Journal of the Caring Sciences*, 7, 113–120.
- Windle, C. (1994). Mental health services research: The shift to research synthesis. *Administration and Policy in Mental Health*, 21(3), 263–267.
- Yin, R. K. (1994). Discovering the future of the case study method in evaluation research. *Evaluation Practice*, 15, 283–290.

The Integration of Basic Research, Treatment Research, and Clinical Practice in Pediatric Psychology

RONALD L. BLOUNT, VICTORIA L. BUNKE,
and JONATHAN F. ZAFF

There are many ways research advances the science of pediatric psychology, as well as clinical psychology in general. Research contributes to the knowledge and thought base on which the future of the field is built. Roberts and McNeal (1995) distinguish among four types of empirical research in pediatric psychology: explicative, assessment, prevention, and treatment research. Each approach contributes uniquely to the study of the clinical phenomenon that are the subject matter of pediatric psychology. However, in recent years the balance among these four types of contributions has become skewed. In particular, there is an overemphasis on explicative research, which examines the associations among variables. In addition, there appears to be a chasm between explicative and treatment research, with explicative research seldom informing the development of treatment programs. However, the chasm has not always been there (see Roberts & McNeal, 1995), and was not characteristic of the early days of the field. Clinical research at that time primarily had an applied goal: to produce clinically significant treatment gains for the patients, as opposed to the current theoretical or model building goal of much of explicative research today. While important, explicative research and the associated theory development were not the primary focus of the majority of research.

RONALD L. BLOUNT and JONATHAN F. ZAFF • Department of Psychology, University of Georgia, Athens, Georgia 30602. VICTORIA L. BUNKE • Department of Educational Psychology, University of Georgia, Athens, Georgia 30602.

The goal of this chapter is to suggest possible ways to reintegrate explicative and treatment research.

This chapter will briefly elaborate on the changing trends in the proportions of explicative and treatment research over the last several decades, probable reasons for these changes, and some of the disadvantages of that trend. Clinical practice will be presented as an arena in which explicative and treatment activities are integrated, with each informing the other. Borrowing from the paradigm of clinical practice, a research program that integrates explicative and treatment research in the area of pediatric pain will then be presented as a model for this approach. Some salient features of the model will be highlighted, along with some additional considerations.

EXPLICATIVE AND TREATMENT RESEARCH: THEY USED TO BE FRIENDS BUT THEY HAVE DRIFTED APART AND DO NOT TALK MUCH ANYMORE

Explanatory research aims to discover associations among variables that may be related to the onset, severity, and course of the clinical phenomenon of interest. In addition, explanatory research seeks to identify those variables that predict the likelihood of developing a disorder; factors associated with better and worse adjustment to, or forms of, a disorder; and factors that relate to the duration of suffering associated with a disorder. The findings from explanatory research may be used to develop theoretical models of how various factors associated with a disorder are related, refine our understanding of the clinical phenomenon of interest and its similarities to and differences from other disorders, and design treatment programs for the alleviation of suffering. Excellent and well-known examples of theoretical models in pediatric psychology based on explanatory research include the risk and resistance model of adaptation to chronic illness and handicapping conditions in children (e.g., Wallander, Varni, Babani, Banis, & Wilcox, 1989) and the transactional stress and coping model for children (Thompson, Gustafson, Hamlett, & Spock, 1992). Explanatory research can provide the data for the inductive process of conceptualizing or theorizing about the phenomenon of interest, with the validity of the theory or conceptualization being verifiable through research designed using deductive processes. In applied areas, the testing of the theory is, in applied areas, accomplished by conducting experimental investigations based upon empirically derived clinical treatment procedures.

In recent years, researchers often have stopped at the level of explanatory research and the associated theoretical model development and refinement, failing to translate their findings into the design of treatment programs. In a survey of articles appearing in the *Journal of Pediatric Psychology*, Roberts (1992) (also see Elkins & Roberts, 1988) found that those studies that were explanatory in nature accounted for approximately 75% of the research published during the period from 1988 to 1992. Assessment research, which was related primarily to diagnosis, testing, surveys, and instrument development and validation, accounted for 12.9%. Prevention was the primary theme in 3.8% of the articles. Only 9.1% of the published research in the journal during that period was intervention or treatment research. Although a large amount of attention is being devoted to explanatory research, it seems that the resulting plethora of explanatory data now rarely informs

the development of clinical treatment programs, as it originally did. We believe that these trends are reflective of the field of clinical psychology as a whole and of the social sciences in general.

There are a number of possible reasons for the tendency of researchers to conduct explicative rather than treatment research. Often with explicative research large databases are gathered and the researcher's attention is focused more on the relationships among variables than on how those variables might be changed in eventual clinical treatment research. From such large databases, numerous publications are often generated, aided by peer reviewers' and journal editors' apparent affinity for theoretical and other forms of explicative research. This high productivity is reinforced through pay raises and the approval of colleagues and granting agency officials. In addition, there is the practical issue of the ease of running additional analyses on an existing large database versus collecting and processing new data, as would be required with treatment research. Working with a large database can be relatively safe in that the probability that some research findings will be published does not depend on the uncertain outcome of a treatment intervention. The advent of sophisticated statistical techniques, such as LISREL and other variations of causal modeling procedures, and even the increased use of multiple regression techniques further promote explicative research. These types of statistical procedures are enticing tools that can provide considerable insights into the associations among variables. However, we suspect that once a person is familiar with the benefits of using a tool, for example, the hammer of causal modeling, many objects, including one's general approach to research, can come to resemble nails.

The biases of many researchers toward conducting explicative research also is supported by anecdotal comments. For example, at a recent presentation a well-known researcher in pediatric and child clinical psychology was questioned about how he had used his extensive explicative research publications and the associated theory development to inform the design of clinical treatments. He replied that he had not, and that he "really did not find treatment research to be very interesting." Indeed, much of explicative research is more cerebral than the pragmatism that characterizes treatment research. However, we believe the mind-set that this researcher's statement reflects does not serve very well the needs of the children and families with whom we work, or the field as a whole. In terms of scientific issues, experimental treatment research provides a very strong test of the validity of the correlational findings derived from explicative research. More importantly, a primary goal in pediatric psychology, as an applied area of clinical psychology, is the prevention or reduction of suffering and the enhancement of well-being. What is the point if the findings from explicative research are not eventually translated into treatment interventions?

Training in the perspective that explicative research is superior to applied, pragmatic, treatment-oriented research begins early in the social scientist's career. In many graduate programs, a lack of theory building or theory testing for a thesis or dissertation often brings negative evaluations by some of the voting faculty members on committees. The threat of failure is enough to assure that most clinical students will attempt to design their study around theory development or testing. Relating research to theory is not negative, but it can be overdone. Committees also frequently call for students to add additional measures. Even though this may

enhance the value of the study, it can promote the tendency to "throw in" yet another measure. Multiple measures are a key feature of large databases used in explicative research. An additional reason for the preponderance of explicative research in graduate programs is the need for the student to finish the research in a short period of time. Large databases, which may already have been collected, perhaps under the direction of the student's major professor, promote the student's design of time-efficient studies using those data. In that situation, analyses need only be conducted from the existing data, as opposed to the student having to run subjects and secure a possible research site. Each step in research takes time for someone who is seeking their degree in a timely manner. A popular maxim is that one of the best predictors of future behavior is past behavior. This is also true for the conduct of primarily explicative research by pediatric psychology and other graduate students, who later become the next generation of researchers in the field (see Blount, Frank, & Smith, 1993, and the entire Special Issue of the *Clinical Psychologist* on training researchers in clinical psychology).

An additional factor which contributes to the lack of treatment research is that many graduate students have not been trained in conducting research in applied pediatric and child clinical settings, where treatment-related issues are very hard to deny. Indeed, if they have clinical experience in hospital and other applied settings, they often encounter psychologists who face extraordinary and ever-increasing demands for clinical services in order to justify their salaries. This nationwide trend threatens to push psychologists into either a clinical or research track. Psychologists in the clinical track have little time for research. In contrast, psychologists in a research track may have little involvement in clinical service provision. This clinician-researcher schism creates a dynamic whereby the quantity and quality of treatment research suffers. In this professional environment the heuristic and compelling aspects of clinical treatment needs have little impact on those who conduct research. In addition, this situation further limits students' access to professional role models who are conducting research evaluating clinical treatments.

Another reason for the bias of researchers toward conducting explicative research could be the relative unattractiveness of some forms of treatment research. In addition to the risk that an intervention does not produce the desired clinically or statistically significant results, in our experience greater effort is usually required for fewer research publications when conducting treatment studies as compared to conducting explicative research. The payoff to the researchers' vita simply is not as likely or as frequent as it is with treatment research. Also, it may be difficult to gain adequate control over the independent variable in much of treatment research. Explanatory research requires only measurement of existing conditions, not changing those conditions in ways that match the intended treatment manipulations.

One danger in not integrating explanatory and treatment research is that the investigations in each area can quickly become stale. For instance, in clinical treatment research and service provision, a prime source of inspiration for developing treatment programs comes from using treatments that have proved successful in other areas. In initial research in a new area, this approach is certainly acceptable. However, this strategy may quickly grow old as study after study shows that treatment X works with yet another problem. This phenomenon is not peculiar just

to treatment research, as there are many examples of a theoretical model and the associated explicative research paradigm being applied to one clinical syndrome or disorder after another. This rehashing is much more likely to occur as the field becomes more mature. There are simply fewer new approaches to be applied in that field. Also, as research about a clinical phenomenon accumulates and the field matures, the initial big treatment gains or leaps in understanding are typically supplemented by contributions that provide incremental treatment effectiveness, greater efficiency of treatment application, comparisons of different treatments, combinations of treatments, or clarification of our understanding of the area. A big leap forward phase in the progression of a field, followed by incremental or refining smaller steps, is characteristic of both treatment and explicative research. Documentation that this is occurring in clinical psychology is indicated by recent efforts to categorize some psychological treatments as having demonstrated effectiveness for particular disorders (Chambless et al., 1996). This categorization of treatments would not have been possible 20 years ago.

Finding a wellspring of creative, effective, efficient approaches or combinations of approaches to the design of treatments then becomes the challenge. Researchers should be open to inspiration from within and from outside of their areas of expertise and their disciplines. One of the most obvious sources of treatment options—the findings from explicative research—is too seldom used. In part we believe this is due to the nature of the explicative research that has been conducted, with a primary focus on distal variables that are not easily manipulated. No doubt there are also many clinically useful findings from explicative research that are not being extended to design treatments. Historically, explicative research was integral to the process of understanding clinical phenomena. Intervening without having a satisfactory understanding of the phenomenon of interest is likely to lead to ineffective, inefficient, or counterproductive treatment outcomes. It will be to the benefit of all involved to get these two old friends talking again.

CLINICAL PRACTICE AS A MODEL OF HOW CLINICAL RESEARCH COULD BE CONDUCTED: BACK TO THE FUTURE?

In clinical practice, understanding the factors that are related to and seem to control the occurrence of a behavior or disorder and the provision of treatment for that disorder are intimately related. For any patient with any problem, providing treatment without a proper understanding of the factors that control the problem is misguided at best and potentially harmful and unethical. The assessment phase of clinical work takes the form of a behavioral assessment funnel (e.g., Cone & Hawkins, 1977; Cronbach, 1970). It is so named because the assessment is initially broad, and with low resolution or focus on specific areas. A number of aspects of a person's life may be assessed prior to effectively defining the problem. Once the problem areas are ascertained, the emphasis narrows and becomes more focused on specific areas. This narrow phase of the funnel, with its high resolution on the target behaviors and the factors that control them, continues throughout treatment. The therapist remains flexible in that the funnel is broadened as necessary when moving to new treatment targets, refining the intervention, assessing for collateral therapeutic changes or side effects, or when the treatment fails (see Table 1).

Table 1. Assessment in the Service of Treatment in Clinical Work

-
- I. Assessment. This phase is equivalent to assessment or explicative types of research. Assessment takes the form of a behavioral assessment funnel (Cone & Hawkins, 1977; Cronbach, 1970), moving from broadband, low resolution (I.A) to narrowband, high resolution (I.B) as the definition of the problem and probable means of intervention emerge.
- A. Mouth of the behavioral assessment funnel:
 1. Information gathered across numerous areas of functioning.
 2. Standardized assessment instruments are often employed.
 3. Numerous people in the patient's life may be contacted.
 4. Patterns and parameters of general behavioral, emotional, intellectual, spiritual, and physiological strengths and weaknesses, as well as resources in the environment, are assessed.
 5. Some initial problem definition and conceptualization emerges.
 - B. Stem of the behavioral assessment funnel. Once the problem behavior(s) is initially defined, a functional analysis is conducted to determine the antecedents and consequences of that behavior.
 1. Included in the consideration of antecedents and consequences are environmental and situational factors, as well as the patient's own emotions, cognitions, physiological states, and overt behaviors that precede, occur during, or follow the problem behavior.
 2. A working theory or conceptualization is formed of the factors that influence the problem behavior.
 3. A formal baseline may be taken of some aspect of the problem behavior and its relevant antecedents and consequences.
 4. Baseline assessments often use very individualized and specific measures.
 5. Baseline assessments may result in a refinement of the conceptualization and subsequent treatment plan.
 6. Assessment using this methodology continues throughout treatment.
- II. Treatment. The treatment is empirically derived based on the assessment and subsequent conceptualization.
- A. Planning of treatment
 1. Probable interventions are considered, including rearranging antecedents, rearranging consequences, training alternative behaviors, or some combination of these three approaches.
 2. Changes in the environment and/or some aspect of the patient are considered as focal points of the intervention.
 - B. Provision of treatment. Continue the narrowband assessments, while maintaining a periodic view of the relevant broadband factors.
 1. The primary method of evaluation is by using the measure employed during baseline.
 2. The clinician also periodically assesses beyond the dependent variable for collateral beneficial or detrimental changes. Adjustments in treatment strategy are made accordingly.
 - C. Evaluation of treatment
 1. Changes in the dependent variable are assessed throughout treatment. These changes provide the basis for continuing to provide treatment as originally planned, revising the specifics of the treatment plan, returning to the broadband or functional analysis stages to reassess and reconceptualize the factors which control the problem behavior and subsequently redesigning the treatment, or selecting new targets for intervention.
 2. Standardized measures also may be employed for assessing treatment effectiveness, but these are typically more related to general functioning than is true of the more specific, more process-related variables that are assessed in the baseline and treatment phases.
 3. The original conceptualization, equivalent to the theoretical model stage in explicative research, is either supported, amended, or discarded based on treatment responsiveness.
-

In the behavior analytic paradigm, the determination of factors that influence the target behavior is referred to as conducting a functional analysis. For the behavior of interest, the antecedents and consequences are ascertained based on interview, observation, and other methods, in order to determine the factors that seem to relate to increased or decreased likelihood of the behavior. Antecedents and consequences, along with the behavior(s) of interest, could be either the patient's own overt behavior, physiological functioning, cognitions, emotional states, or external stimuli. Antecedents are events that precede the behavior and increase the likelihood of the behavior occurring. Consequences are those events that follow the behavior and serve either to increase or decrease the likelihood of the behavior occurring. These functional antecedents and consequences tend to occur in close temporal proximity to the target behavior, but temporally distal antecedents and consequences may be the controlling factors or at least exert some control. However, distal factors are usually much harder to manipulate and their impact on behavior tends to be less direct than the temporally proximal factors.

After initial data gathering, the clinician analyzes this information and forms a conceptualization of the factors that control the behavior of interest. This phase is equivalent to conducting the correlational analyses and associated conceptual or theory development phase that characterizes explicative research. For the clinician to consistently stop at the level of explanation, understanding, and the associated conceptualization or theory development would be unthinkable and unacceptable. In like manner, clinical researchers also should move beyond this point.

Following the functional analysis, the selection of the best route of intervention is made. The intervention could include rearranging antecedents, teaching new behaviors, rearranging consequences, or some combination of the above. Treatment is then initiated, with careful observation of the behaviors of interest and other concomitant behavioral changes, whether beneficial or detrimental. Efforts are made to assure that the independent variable is manipulated in the desired manner so as to better differentiate between true treatment failures and simple patient noncompliance. The prescription for what the therapist should do in the case of the treatment not being implemented properly, or a true treatment failure, would of course differ in these two situations. The goal is to produce clinically significant changes in behavior. Statistical significance is of little value for the individual patient. Ongoing monitoring is essential, with each occasion of assessment being an opportunity to confirm, deny, or amend the original theory or conceptualization of controlling antecedents and consequences and how to modify them. This provides a means of experimentally validating the conceptualization or theory of how factors influence the patient's behavior. If the results are turning out as desired, the theory tends to be confirmed. If not, it is either revised or discarded altogether.

In clinical practice, at least as modeled after applied behavior analytic thought, assessment and treatment are closely connected at all points. Assessment informs treatment and treatment responsiveness or outcome informs additional assessment and possible theory or conceptual revisions. In terms of scientific methodology, this latter phase is analogous to experimentally testing and validating the results of explicative research. The revised theory is then tested and the process begins again. The applied behavior analytic therapist–researcher talks about “coming under the control of his or her data,” meaning that he or she is responsive to the patient—

subject, and adjusts theory and approach accordingly. There is a high degree of focus and resolution on the behavior of the individual patient during the assessment (explicative) and treatment stages. At each stage, the data gathering, conceptualization, and treatment are pragmatic in that they are in the service of producing clinically significant changes in behavior. Some information that is gathered may not prove useful for facilitating behavior change, but that is inevitable. However, merely changing the theory or refining the explanation of how the variables interrelate is not the goal, nor should it be. The approach described, along with much elaboration, is the hallmark of applied behavior analysis and is widely applicable to both clinical treatment and clinical treatment research. For a further explanation of some of the basic tenets of this approach, the reader is referred to the classic paper by Baer, Wolf, and Risley (1968). It is our experience that the basic tenets of applied behavior analysis are not very much practiced by clinicians today, even in behavioral circles, and much less so outside of behavioral circles. However, we believe that some of the basic principles of applied behavior analysis could be employed, perhaps regardless of the theoretical paradigm of the researcher, and that this employment would go far to encourage the development of empirically derived treatment programs.

ACUTE PROCEDURAL PAIN IN CHILDREN: A MODEL FOR THE INTEGRATION OF EXPLICATIVE AND TREATMENT RESEARCH

As an example of integrating explicative and treatment research, we will focus on our work with acute procedural pain in pediatric patients. The process and the thinking by the researcher that went on behind the pages of the published research are what we wish to share.

Explanative Research: Begin with the End in Sight

While following the families on a consultation-liaison service, I (RLB) was blessed to observe some important information about pediatric oncology patients, their parents, and the staff during the children's lumbar punctures (LP) and bone marrow aspirations (BMA). During these observations, it soon became apparent that some children displayed less distress and more coping behaviors than others. More importantly, in terms of therapeutic implications, the parents and medical staff interacted with the children in different ways. Some of these ways seemed to encourage coping, while other ways seemed to encourage distress.

My initial interest in children with cancer undergoing BMAs and LPs was clinical, with numerous preliminary observations being made. We asked frequent questions of the staff and examined the literature. Vicki Wolfe, who had studied marital interactions, suggested using a technique called *sequential analysis* as a means of mapping the flow of the interactions that occurred in the treatment room. This method allows for the determination of temporal antecedents and consequences of particular child and adult behaviors. The discovery of temporal antecedents and consequences could pave the way for the determination of important functional relationships. Being on a very low budget at the time, we set about audiotaping the children and others present prior to, during, and after the medical

treatments. We later devised an assessment instrument for coding the interactions (Blount et al., 1997). None of the children and parents had been trained in the use of coping strategies.

This explicative research was designed to learn more about what happens in the medical treatment room. The research was patterned after the functional analysis that is used in clinical service provision, in which one of the authors (RLB) was trained. However, it differed from clinical practice in that this functional analysis investigated the antecedents and consequences of child coping and distress behavior, as well as the other child-adult and adult-adult interactions, of the group of children, parents, and staff. Hence, we were interested in learning about generalizable findings. Three investigations were conducted using the same data.

In our initial research in this area (Blount et al., 1989), using a sequential analysis framework and the 35 code Child-Adult Medical Procedure Interaction Scale (CAMPIS), we determined those adult behaviors that typically preceded or followed children's distress and coping behaviors. We found that the child coping behaviors of nonprocedural talk and humor by the child (forms of distraction) were typically preceded and followed by nonprocedural talk and humor directed to the child by the medical staff and/or parents. Also, the child coping behavior of use of deep breathing seldom occurred unless it was preceded by adults repeatedly prompting the children to breathe. This suggested that frequent adult prompts were necessary in order to promote the occurrence of the desired coping behaviors.

In contrast, child distress behaviors were typically preceded and/or followed by adult's reassuring comments, empathic comments, apologies, criticism, and giving the child control over the beginning of some aspect of the medical procedure. We termed these behaviors *distress promoting*. By far, reassurance by parents and staff was the most commonly occurring of the behaviors to precede and follow child distress. While this finding did not indicate a causal relationship, it did suggest that reassurance, empathic comments, apologies, criticism, and giving control to the child were much more related to the occurrence of distress behavior than to coping behavior. The findings of this study had direct implications for the design of therapeutic programs. For example, in order to promote child coping, it would be necessary to train the children to use the coping behaviors indicated by the explicative findings, and train the adults to repeatedly prompt their occurrence using the appropriate cues. Child distress was also associated with specific adult antecedents. However, because increases in child coping and adults' coaching of their children to cope were incompatible with child distress and adult distress-promoting behaviors, respectively, we elected not to target these undesired behaviors directly in the first series of treatment studies.

Our subsequent analyses of the data (Blount, Sturges, & Powers, 1990) indicated that there were phase variations in the coping behaviors that children used and similarly in the adults' prompts that facilitated those coping behaviors. During the anticipatory phase prior to the beginning of the painful aspects of the medical treatment, the most common coping behavior was nonprocedural talk and occasional humor by the child. Later, during the painful medical treatment, those coping behaviors seldom occurred and the child instead shifted to the use of deep breathing. These types of child coping behaviors were closely associated with the adult's use of nonprocedural talk and humor to the child and with coaching them to breathe, which occurred in the same phase-specific manner. We did not find a

general child or adult coping-promoting disposition. That is, use of nonprocedural talk by or to the child in the anticipatory phases was not correlated with breathing or coaching the child to breathe, respectively, in the painful phases. Child distress and the associated adult behaviors were highly correlated from the anticipatory to the painful phase, suggesting that once a child started on a chain of distress behaviors, it was difficult to break. These data refined the treatment implications of the previous study, suggesting that the particular phases had particular coping demands, and that parents and staff needed to vary their prompts to the children across phases in order to promote the desired coping behavior. In order to reproduce the coping pattern found in these untrained subjects, the findings indicated that children should be taught to use one type of coping behavior prior to the medical procedure and another during the painful aspects. Adult's prompts would have to vary similarly.

In the final analyses of the initial data (Blount, Landolf-Fritzsche, Powers, & Sturges, 1991b), we rank-ordered the children on the basis of their use of coping behaviors. We found that parental prompts varied as we expected, with the high-coping children being more likely to have parents and staff prompt them to cope, while doing fewer of those behaviors previously found to be associated with distress. Also, the high-coping children were found to be more responsive to the prompts by parents and staff to engage in coping behaviors. However, for both the high- and low-coping children, the same adult-child behavioral relationships were found. Both high- and low-coping children were more likely to cope when coping-promoting prompts were present and to show distress when reassurance, empathy, apologies, criticism, and giving control to the child occurred. This suggested generalizability of the findings from the previous analyses to children with different levels of coping. In addition, there were no differences found in children's reactions to various prompts from a parent versus a staff member. Again this suggests generalizability of the interactional pattern of behavior between the child and either the parent or the medical staff member.

Another finding of considerable importance was the behavioral associations among the adults' behaviors in the previous studies, particularly the first study (Blount et al., 1989). In that investigation, the most common behavior to precede and follow most of the adult behavioral categories was another adult behavioral category of the same type. What this suggested was that adults in the treatment room take many of their cues from each other as to how to behave toward each other as well as toward the child. In terms of therapeutic implications, it suggests that it may not be necessary to change the behavior of all the people in the treatment room directly. Sufficient changes in the behavior of one of the participants in the medical treatment room may be adequate to change the behavior of all other participants.

Based on our explicative research and the work of others we developed a working model or conceptualization of how various factors impact coping and distress behavior in children undergoing painful medical procedures (Fig. 1). There are primarily two categories of independent variables: those that are temporally and functionally proximal and those that are temporally and functionally distal to the behaviors of interest. The behaviors of interest are the coping and distress behaviors of the child that occur before, during, and after the medical procedure. The proximal behaviors, which are more situational or statelike, are the parental and staff in-session behaviors that occur in the treatment room and exert a powerful

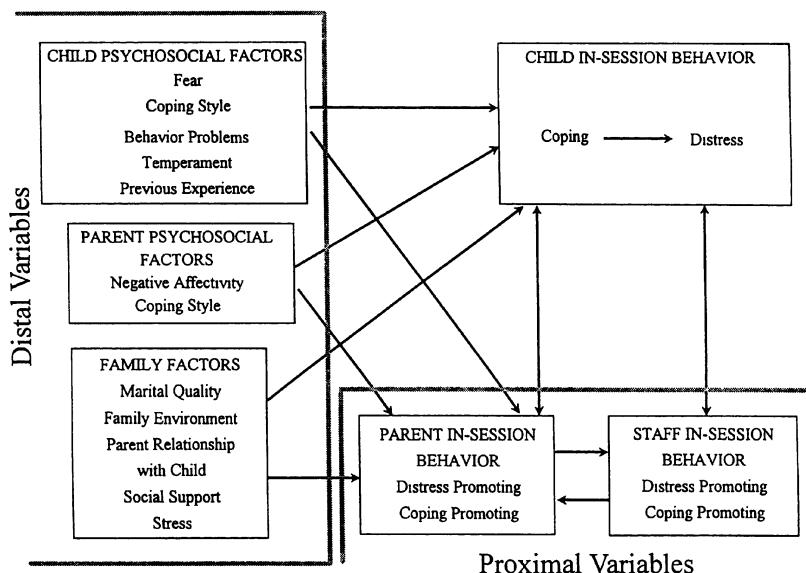


Figure 1. The Proximal–Distal Model of Children’s Coping and Distress During Painful Medical Procedures. This model was changed slightly from the one presented by Varni, Blount, Waldron, and Smith (1995).

influence on the child’s reactions. For example, in a recent study of preschool children receiving immunizations who were not trained in coping behaviors, parent and staff in-session behavior accounted for 38% of the variance in child coping and 55% of the variance in child distress (Frank, Blount, Smith, Manimala & Martin, 1995). As described above, these proximal behaviors should vary in a phase-specific manner with the unique coping demands of the different parts of the medical procedure. The proximal variables also tend to be the more easily changed of the two categories of variables, at least for the behaviors of child coping and distress. These proximal variables occur in the narrow-focus, high-resolution part of the behavioral assessment funnel described earlier.

In contrast to the proximal variables, distal variables occupy the broad-band, low-resolution part of the behavioral assessment funnel. Some distal variables, such as the child’s age, fear, and level of distress during past medical treatments, are significantly correlated with the child’s distress (e.g., Blount, Davis, Powers, & Roberts, 1991a; Dahlquist et al., 1986; Frank et al., 1995; Jay, Ozolins, Elliott, & Caldwell, 1983; Pate, Smith, Blount, & Cohen, 1996). However, these variables may be difficult or impossible to change or, in the case of fear, may be changed in the process of teaching coping behaviors. The usefulness of these distal variables seems to be as a screening mechanism for identifying those characteristics of patients who may be in need of training or for serving as a marker for the presence of potentially modifiable, functional, proximal variables. Many of the other more distal variables seem to have little direct impact on child behavior during acute

painful medical procedures. As opposed to the situational proximal variables, distal variables tend to be more trait-like.

The Goal: Empirically Derived Treatment Research

A unique advantage of conducting empirically derived treatment research is that it provides a powerful test of the validity of the findings from the explicative research. As such, it either validates, disproves, or, almost always, refines the original conceptualizations that were associated with the explicative research. Our own treatment research certainly forced some refinement and adaptation of our original conceptualizations. Explicative research that examines the associations among variables that are not easily manipulated, typically the more temporally and functionally distal variables, lacks this degree of acidity in any test of the theory or conceptualization.

The initial three treatment studies conducted by our group incorporated a matching-to-sample approach. The desired sample was the high-coping children, as well as the coping/coping-promoting, phase-specific interaction patterns found in our explicative research. Our goal was to promote the same sort of adult-child interactions in the subjects in the treatment studies which characterized the high-coping children during the various medical phases in our explicative research. In the first treatment study conducted (Blount, Powers, Cotter, Swan, & Free, 1994), we attempted to teach children and parents to discuss nonprocedural topics and to deep breathe in a manner consistent with the behavior of the children in our explicative research. Both settings were pediatric oncology treatment centers, although at different hospitals, and the children in the explicative research were a few years older. We were unable to successfully teach the children and parents to use the same nonprocedural talk or breathing behaviors at a comparable level that the children in the explicative research had used. In retrospect, in the setting in which the explicative research was conducted, it is likely that one laboratory technician and one nurse present during the procedures were the driving forces behind the parents' and other medical staffs' use of so many prompts for the children to cope. We have never observed so much "naturally occurring" deep breathing or distraction by children undergoing any painful medical treatments in any other setting.

As a fallback technique, we kept the same concepts and used props to facilitate distraction and breathing by the children. In our experience, this type of flexibility is necessary when moving from explicative to treatment research. The props that facilitated adults' coaching of the children were various toys that could be used before the medical procedures began. Incorporating toys helped eliminate the silence that would otherwise often ensue during the many minutes prior to the medical treatment. The prop that facilitated breathing in these young children was a party blower. Selecting this item for breathing promotion was based in large part on a conversation with Bill Redd, who was then associated with Memorial Sloan-Kettering, along with his collaborators, Sharon Manne and Paul Jacobsen. The party blower also had a fairly heavy distraction component. Use of the party blower was easily verifiable, thus assuring that the desired coping response was occurring and allowing for a better determination of the supposed inverse relationship between coping and various forms of child distress and pain. Training or booster

training sessions, which incorporated role play, rehearsal, and feedback, were conducted prior to each LP except the last one. In a multiple baseline across subjects design, the treatment worked very well for two of the subjects who adapted to using the coping behaviors regularly. One of these two children and their parents immediately adapted to the use of the desired coping behavior. The other, a younger child, took several sessions before her level of blower usage was adequate. Therefore, the necessity of training until skill acquisition was apparent, as was also indicated for the next subject. The third parent-child dyad used distraction and book reading throughout the first treatment session, with little to no reliance on a blower during the LP. Distress was extremely low compared to baseline. However, in subsequent sessions, the child's distress was at or above baseline levels. Heightened family disruption between the first and subsequent treatment sessions was the suspected culprit for this setback in coping and coping-promoting ability.

We successfully replicated this basic training procedure in another study using pediatric oncology patients undergoing intramuscular (IM) and intravenous (IV) injections, again using a multiple baseline design (Powers, Blount, Bachanas, Cotter, & Swan, 1993). Based on parent and child preference, the use of parent-prompted counting was substituted for use of the party blower for some of the children in the study. Therapeutic effects were obtained and maintenance was found for these subjects.

In the final study in this series of treatment investigations (Blount et al., 1992), we took what we had learned from experimentally controlled single-subject designs with pediatric oncology patients who underwent repeated painful procedures and applied the basic training program to 4- to 6-year-old preschool children undergoing immunizations at a county health department. Due to the high volume of patients seen at the health department and the nonrecurring nature of this injection compared to recurring LPs, the training period was shortened to 7–10 minutes. We eliminated the requirement to demonstrate the desired skills to proficiency prior to ending a training session, as was used in the first two studies. The investigators presented the training program using role play and encouraged the parent and child to practice while receiving feedback. This represented a move toward broader application of the training techniques used in the previous investigations. Training resulted in therapeutic gains and less distress on some but not all dependent variables. In addition, we found that the untrained nurses also took their cues from the trained parents and coached the trained children to cope. This latter result also supported the findings from our explicative research and confirmed that not all people in the treatment room need to be trained in order to assist in coping-promoting efforts.

In one of the most easily used techniques we know for fostering children's coping (Cohen, Blount, & Panopoulos, 1997), we incorporated the concept of adult-prompted distraction of the children's attention from the fearful or painful aspects of the injection to highly appealing cartoons that were watched before, during, and after the injections. Three groups were used, including an attention control, nurse-directed distraction, and nurse-directed distraction plus training children and parents condition. This study was designed to determine whether costly and time-consuming training of each parent and child dyad was necessary, given the compelling distraction of appealing cartoons and nurse prompting to attend. Participants were 4- to 6-year-old children receiving immunizations at a health department.

Upon the child's entrance into the immunization room, nurses offered treatment subjects a choice of viewing *Aladdin*, *Beauty and the Beast*, *Barney*, or *The Lion King*. For the subjects in the treatment conditions, immunizations proceeded as normal except at procedural junctures, such as cleaning for the injection and just before needle insertion, at which time the nurse would direct the child's attention to the video. Nurses also directed the child's attention to the cartoon at signs of child distress. The results indicated that nurse-directed distraction with untrained children and parents was just as effective and less costly than nurse direction plus training the child and parents condition. Both interventions were superior to the no-treatment control condition on all observational, child self-report, parent report, and nurse report measures of pain, distress, and coping. Similar to the previous investigations, untrained parents also joined in and prompted the child to attend to the cartoons, apparently taking their cues from the nurses as to how to interact with their children. As in our previous research (Blount et al., 1992), parents in the intervention conditions were also less distressed than those in the control group. The nurse-directed condition was also judged by the nurses to be easy to use. Because of its cost-effectiveness and the nurses' indication of their intention to continue to use the intervention, this study could be viewed as a move toward treatment dissemination (Blount, 1987). However, to our knowledge, this cost-effective approach has not been widely instituted or adopted as standard health care in other settings. Therefore, in our research and the field as a whole, treatment dissemination is an issue in need of attention (Blount, 1987).

We have conducted one investigation in our laboratory in order to evaluate the presumed negative, causal impact of reassurance on child distress (Manimala, Blount, & Cohen, in press). This study was conducted last because it related more to conceptual-theoretical issues than to practical clinical issues. In our explicative and conceptual work, we described reassurance as a "distress-promoting" behavior. We have never taught parents or staff to engage in less reassurance. Instead, we have always taught them to engage in more distracting and coaching, believing that reassurance would decrease as distracting and coaching increased. However, when referring to the effects of reassurance in our proximal-distal model of variables that influence child distress and coping, as well as in our grouping of the 35 CAMPIS codes into the six code CAMPIS-R (Blount et al., 1997), we use the term *distress promoting*, rather than *distress correlated*, thereby indicating causal associations. Three groups of preschool children undergoing immunizations at a health department served as subjects. There was a control, reassurance, and distraction-coping group. Due primarily to factors peculiar to that busy health department, it was more difficult to gain adequate control over the independent variables. However, all results that were significant were in the predicted direction, with the reassurance subjects showing more distress. Most striking, 40% of the children in the reassurance group required restraint, whereas only 15% of the children in the control and 10% of the children in the distraction group required restraint. This further supports the findings from a different aspect of our explicative research.

An additional point worth noting is that our medical colleagues have been much more interested in treatment research, which has the potential to assist their patients, than they have been in explicative research, which is primarily related to theory building or testing. Individual patients in explicative research often do not benefit much from their participation in those investigations. However, even with

beneficial treatment research, procedures that require significant changes in the way medical regimens are provided or in the medical professionals' interactions with patients may meet with some resistance by the frontline medical personnel. Acceptance of clinical lore, the effort required to adopt new routines, and the demands of clinical service needs, among other factors, all may conspire to prevent change in medical protocol.

HOW TO MAKE THE INTEGRATION WORK: SALIENT FEATURES AND ADDITIONAL THOUGHTS

We have described one example of the successful integration of explicative and treatment research. What are some of the foundations for this successful integration that may be emulated by others? For one, explicative research can be conducted in such a way that it is more likely to produce results that have potential to inform the design of treatment programs. But this requires the explicative researcher to choose his or her variables well for that purpose. From our presentation on conducting a functional analysis in clinical service provision and our review of the explicative research conducted in our laboratory, it is clear that those independent variables that are temporally and functionally proximal to the behaviors of interest are the ones that have the greatest likelihood of exerting an influence on the dependent variable. Discovering the specific details about those functional associations, as was described in the review of our explicative research, should therefore be the goal! As noted throughout this chapter, proximal variables also tend to be the ones that are likely to be most easily manipulated in a training program. However, the degree to which this recommendation can be applied probably varies with the discreteness of the dependent variable, or perhaps more correctly the discreteness with which it is operationally defined. In our area of research, coping and distress are specific behaviors which occur during a painful medical procedure. This medical procedure has an anticipatory, encounter, and recovery phase which starts and ends in a short period of time before and after the predictable event of the injection(s). Other examples of discrete behaviors include an instance of an argument or aggression on the playground, having a stomachache as a way of staying home from school, giving a speech in class, and putting your tray away after lunch. Other situations are not so discrete, such as "adjustment" to chronic illness, "adjustment" to hospitalization, or "adjustment" to parental discord. The stressors and the distress or coping responses in these cases tend to be more prolonged or chronic and multicomponent.

Going back to our model of clinical work, when the clinician encounters a multicomponent task or long-term situation, which is causing the patient difficulty, one possible way to address the problem is to conduct a task analysis and break the complex situation down into more discrete component parts. For example, there are different events within the relatively long-term, complex stressor of hospitalization. Those events may include checking in, putting on hospital clothes, answering strange questions, anesthesia induction, injections, and people coming in and out of your room to wake you so you can take a pain pill. Each of these situations could be rated for unpleasantness or distress. Examples of successful "adaptation" to each of the more discrete situations could be ascertained and the factors that

promoted that adaptation could be determined. In clinical work, this would again be equivalent to conducting a functional analysis for each of the individual situations. It is possible that some overall coping or coping-promoting pattern could be found in this explicative investigation of adjustment to hospitalization. If so, there may be a need to train only one or a few coping or coping promoting behaviors. After the explicative research, creative treatment solutions could be designed and instituted. Overall adjustment should still be assessed, but that is likely to be a product of successfully handling the numerous smaller situations. In fact, the proposed hypothetical approach is similar to the treatment approach used by Visintanier and Wolfer (1975) in their early studies employing stress-point nursing to help children and their families adjust to different aspects of hospitalization.

A variation of the approach that guided stress-point nursing studies was used by Alexandra Quittner and her colleagues (e.g., DiGirolamo, Quittner, Ackerman, & Stevens, 1997; Quittner et al., 1996). As a means of selecting targets for intervention from among the difficulties faced by children and adolescents with cystic fibrosis (CF), these researchers assessed aspects of both the frequency and the difficulty of various problems individuals with CF may encounter. Adolescents with CF then indicated the coping responses they would make to audiotaped vignettes portraying the most frequent and difficult situations. The competence of these coping responses across different situations was then rated by teens, parents, and health care experts. The product of this assessment research is a method that has promise for designing treatment interventions for individual children based on profiles of their coping competency in various domains. Treatment research using this personalized assessment approach is currently being conducted. In addition, using this same methodology, different versions of the coping competency profiles are also being developed for schoolage children and for parents.

Another technique, of which some researchers may not be aware, for using their explicative data to inform the development of treatment is by the use of the matching-to-sample approach. This approach has been used in social skills training, among other areas, and involves teaching those who need treatment to do much the same thing as those who are dealing successfully with a situation. Matching-to-sample does not presume that the researcher necessarily knows what the best treatment is, but lets the data be a primary guide for the design of the treatment. Of course, with this approach the more the people and situation to be trained resemble those who served as the sample, the more likely the treatment will be useful. Issues of generalizability to subjects of different ages, socioeconomic status, ethnicity, and other potentially important dimensions are therefore important to consider.

Because of such within-group differences, we consider it of paramount importance to get to know your subject population. Too often, researchers do not have a good idea of the behavior-patient(s)-situation-environment they are studying before beginning the research. This is the type of format that is most likely to lead to explicative research where the main goal is to relate variables to variables. When it comes to treatment, trying to implement a program without a good understanding of the problem would probably be ineffective or even counterproductive. However, we have heard laypeople and imminent psychologists suggest this "blind" approach. Getting to know the subject population for pediatric psychologists means talking to the patients, parents, medical staff and other professionals involved; learning something about the disease or psychological condition; learning some-

thing about the medical event(s) or stressor(s); coming to know the setting and the routine in that setting; and conducting numerous observations.

We believe direct observation is an extremely important method for understanding the population being studied. For example, in our work with acute distress, we have had well-meaning medical personnel enthusiastically tell us that children experienced high levels of distress with procedure *X*, and that there were many patients seen in a particular clinic. Observation and an examination of the patient contact records may or may not confirm the honest but possibly incorrect impressions (see also Wolfe, Blount, Saylor, Dufour, & Saylor, 1987). Going back to the example of clinical work, conducting a good functional analysis of the presenting problem and the antecedents and consequences that influence it, as well as the setting factors, is essential for good clinical treatment. Being involved in clinical work in the area of research interests is one way to accomplish this. We also like observing and interacting with the subjects because they do so many things to teach us and can provide a source of creative ideas of what to do and what not to do.

This same philosophy of getting to know your subject matter and coming under the control of your data also has guided our transitions from explicative to treatment research. Our first two treatment studies following the explicative research employed multiple baseline experimental designs. These low numbers of subjects (*N*), high-resolution (focus), multiple measurement designs are in contrast to the high number of subjects, low resolution, and often only one measurement occasion that characterizes group designs. The high degree of resolution on both the application of the independent variable (treatment) and the evaluation of the effects of that treatment, allow the researcher to be more sure of how his or her treatment will work or will need to be modified or perfected when moving from explicative to treatment research. Single-subject designs are powerful experimental designs for evaluating the effects of treatment, not just pilot work for group research, although they can function that way too (see Collins, Baer, & Blount, 1985; Gelfand & Hartmann, 1984). Unfortunately, certain single-subject designs are not applicable in all situations. For example, if the behavior happens only once, repeated measurement occasions are not an option. However, clinical service provision or single-subject research to test the application and effects of treatment prior to larger-scale group research seems prudent in the majority of cases.

Another theme that has characterized our research in acute pain is to start with a fine-tooth comb. In our work with acute distress, we started with a more involved, costly, and high-resolution approach, and later moved to less sustained involvement with our subjects, less cost, and lower resolution (see Blount, 1987, for an elaboration on these themes). For example, our CAMPIS coding system is a 35-code system, with 16 codes for child distress, coping, and additional child behaviors. The remaining 19 codes were of adult behaviors directed toward the child or other adults. The speaker was coded for every vocalization. If a child, mother, father, nurse, physician, and another adult were present in the treatment room, that was a total of 111 possible speaker-content combinations. Those speaker-content combinations could occur over nine different phases in our initial research (e.g., Blount et al., 1990). That involved much work, time and effort, but it allowed for recombinations of those content codes or speakers in a variety of ways. If you are conducting a true functional analysis in clinical work or are using an inductive approach in explicative research, you are looking for what the data tell you rather than presum-

ing the data are ready to recite a particular chapter and verse because that is what the preexisting theory or even clinical lore dictates. The researcher can always collapse code categories but can never expand them without recoding of the data.

From our 35-code CAMPIS, we later combined the codes into the six-code CAMPIS-R, which includes children's coping, distress and neutral behaviors and adults' coping promoting, distress promoting, and neutral behaviors (Blount et al., 1997). This recombination was based on conceptual factors, other's research (e.g., Bush, Melamed, Sheras, & Greenbaum, 1986; Jay et al., 1983; Katz, Kellerman, & Siegel, 1980; LaBaron & Zeltzer, 1984), and primarily our own earlier explicative research (Blount et al., 1989). The validity of these CAMPIS-R groupings was supported by a concurrent validity study (Blount et al., 1997), as well as by being experimentally validated by our treatment research. We are now converting the CAMPIS-R into a more easily used rating scale, which still maintains the conceptual advantages of the CAMPIS-R over previous scales in this area. At the time of the writing of this chapter, the validity of this rating scale seems to be good.

The CAMPIS, which takes time and effort to use, was being developed and used in our laboratory at a time when a number of leaders in the field were calling for easily used scales that required less time and effort. We have argued elsewhere (Blount, 1987) that researchers can move prematurely to low-cost assessments or treatments at the sacrifice of effectiveness. Again, we believe that it is better to start with higher resolution and if necessary higher cost, in assessment or treatment, in order to gain a thorough understanding of the subject matter and in the case of treatment to help assure effectiveness. When an assessment instrument or a treatment protocol accomplishes what it is designed to accomplish at as low a cost as possible, only then is it efficient for day-to-day use. Assessment instruments or treatments are not efficient if they are not effective, regardless of the low cost. Another point related to efficiency that is beyond the scope of this chapter is the dissemination of existing cost-effective therapeutic programs. Dissemination is an issue that receives almost no attention in the literature, but one that researchers in any applied health profession should consider (Blount, 1987).

This chapter has been about the process of moving from explicative to treatment research. However, this presupposes that the researcher has determined in his or her mind that conducting treatment research is a worthy activity and that preventing or reducing human suffering is the ultimate goal of psychological investigations. Obviously not all of any active researcher's work is going to be in the service of alleviating suffering, and that is okay. However, clinical psychologists should be thrilled to conduct research that potentially benefits patients and informs those who provide frontline service in schools, clinics, and health care settings.

Finally, what we have proposed is only one way that researchers can move from explicative to treatment research. There are probably a number of variations on the theme we described (e.g., Quittner et al., 1996), as well as novel approaches. This chapter was based on many subjective impressions, personal perspectives and experiences, and personal biases. Others may not even consider the approximate 7.5:1 ratio of explicative to treatment research in the *Journal of Pediatric Psychology* (Roberts, 1992), and probably the field of clinical psychology and the social sciences as a whole, to be a problem. However, we and at least some others believe that the explicative-to-treatment research ratio can and should be reduced through

various means, including following the guidelines proposed in this chapter. Whether these approaches will be adopted will depend on not just the individual researcher's decision, but also on the research environment in graduate programs, journal editors' and reviewers' decisions as to what gets published, granting agencies deliberations about what gets funded, the degree of entrenchment of social scientists' belief that theoretical research is the acme of research investigations, and administrators' commitment to the conduct of treatment research as having a unique value. One chapter clearly will not suffice to accomplish all of that, but it is a start, and one that needs to be made.

ACKNOWLEDGMENT. We wish to thank Sandy M. Bouldin for her helpful comments on an earlier version of this chapter.

REFERENCES

- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91–97.
- Blount, R. L. (1987). The dissemination of cost-effective psychosocial programs for children in health care settings. *Children's Health Care*, 15, 206–213.
- Blount, R. L., Bachanas, P. J., Powers, S. W., Cotter, M. C., Franklin, A., Chaplin, W., Mayfield, J., Henderson, M. & Blount, S. D. (1992). Training children to cope and parents to coach them during routine immunizations: Effects on child, parent and staff behaviors. *Behavior Therapy*, 23, 689–705.
- Blount, R. L., Cohen, L. L., Frank, N. C., Bachanas, P. J., Smith, A. J., Manimala, M. R., & Pate, J. T. (1997). The Child–Adult Medical Procedure Interaction Scale—Revised: An assessment of validity. *Journal of Pediatric Psychology*, 22(1), 73–88.
- Blount, R. L., Corbin, S. M., Sturges, J. W., Wolfe, V. V., Prater, J. M., & James, L. D. (1989). The relationship between adults behavior and child coping and distress during BMA/LP procedures: A sequential analysis. *Behavior Therapy*, 20, 585–601.
- Blount, R. L., Davis, N., Powers, S., & Roberts, M.C. (1991a). The influence of environmental factors and coping style on children's coping and distress. *Clinical Psychology Review*, 11, 93–116.
- Blount, R. L., Frank, N. C., & Smith, A. J. (1993). Training the next generation of researchers in clinical psychology. *Clinical Psychologist*, 46, 100–105.
- Blount, R. L., Landolf-Fritsche, B., Powers, S. W., & Sturges, J. W. (1991b). Differences between high and low coping children and between parent and staff behaviors during painful medical procedures. *Journal of Pediatric Psychology*, 16, 795–809.
- Blount, R. L., Powers, S. W., Cotter, M. W., Swan S. C., & Free, K. (1994). Making the system work: Training pediatric oncology patients to cope and their parents to coach them during BMA/LP procedures. *Behavior Modification*, 18, 6–31.
- Blount, R. L., Sturges, J. W., & Powers, S. W. (1990). Analysis of child and adult behavioral variations by phase of medical procedure. *Behavior Therapy*, 21, 33–48.
- Bush, J. P., Melamed, B. G., Sheras, P. L., & Greenbaum, P. E. (1986). Mother-child patterns of coping with anticipatory medical stress. *Health Psychology*, 5, 137–157.
- Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Srits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *Clinical Psychologist*, 49(2), 5–18.
- Cohen, L. L., Blount, R. L., & Panopoulos, G. (1997). Nurse coaching and cartoon distraction: An effective and practical intervention to reduce child, parent, and nurse distress during immunizations. *Journal of Pediatric Psychology*, 22(3), 355–370.
- Collins, F. L., Baer, R. A., & Blount, R. L. (1985). Single-subject research designs for optometry. *American Journal of Optometry and Physiological Optics*, 62, 516–522.
- Cone, J. D., & Hawkins, R. P. (1977). Introduction. In J. D. Cone & R. P. Hawkins (Eds.), *Behavioral assessment: New directions in clinical psychology* (pp. xiii–xxiv). New York: Brunner/Mazel.

- Cronbach, L. J. (1970). *Essentials of psychological testing* (3rd ed.). New York: Harper & Row.
- Dahlquist, L. M., Gil, K. M., Armstrong, F. D., DeLawyler, D. D., Greene, P., & Wouri, D. (1986). Preparing children for medical examinations: The importance of previous medical experience. *Health Psychology*, 5, 249-259.
- DiGirolamo, A. M., Quittner, A. L., Ackerman, V., & Stevens, J. (1997). Identification and assessment of ongoing stressors in adolescents with a chronic illness: An application of the behavior analytic model. *Journal of Clinical and Child Psychology*, 26, 53-66.
- Elkins, P. D., & Roberts, M. C. (1988). *Journal of Pediatric Psychology*: A content analysis of articles over its first 10 years. *Journal of Pediatric Psychology*, 13, 575-594.
- Frank, N. C., Blount, R. L., Smith, A. J., Manimala, M. R., & Martin, J. K. (1995). Parent and staff behavior, and previous child medical experience, and maternal anxiety as they relate to child distress and coping. *Journal of Pediatric Psychology*, 20, 277-289.
- Gelfand, D. M., & Hartmann, D. P. (1984). *Child behavior analysis and therapy* (2nd ed.). New York: Pergamon.
- Jay, S. M., Ozolins, M., Elliott, C. H., & Caldwell, S. (1983). Assessment of children's distress during painful medical procedures. *Health Psychology*, 2, 133-147.
- Katz, E. R., Kellerman, J., & Siegel, S. E. (1980). Behavioral distress in children undergoing medical procedures: Developmental considerations. *Journal of Consulting and Clinical Psychology*, 48, 356-365.
- LaBaron, S., & Zeltzer, L. (1984). Assessment of acute pain and anxiety in children and adolescents by self-reports, observer reports, and a behavior checklist. *Journal of Consulting and Clinical Psychology*, 52, 729-738.
- Manimala, M. R., Blount, R. L., & Cohen, L. L. (in press). The effects of parental reassurance versus distraction on child distress and coping during immunizations. *Children's Health Care*.
- Pate, J. T., Smith, A. J., Blount, R. L., & Cohen, L. L. (1996). Childhood medical experience and temperament as predictors of adult functioning in medical situations. *Children's Health Care*, 25, 281-296.
- Powers, S.W., Blount, R.L., Bachanas, P.J., Cotter, M.C., & Swan, S. C. (1993). Helping preschool leukemia patients and their parents cope during injections. *Journal of Pediatric Psychology*, 18, 681-695.
- Quittner, A. L., Tolbert, V. E., Regoli, M. J., Orenstein, D., Hollingworth, J. L., & Eigen, H. (1996). Development of the Role-Play Inventory of Situations and Coping Strategies (RISCs) for parents of children with cystic fibrosis. *Journal of Pediatric Psychology*, 21, 209-235.
- Roberts, M. C. (1992). Vale dictum: An editor's view of the field of pediatric psychology and its journal. *Journal of Pediatric Psychology*, 11, 785-805.
- Roberts, M. C., & McNeal, R. E. (1995). Historical and conceptual foundations of pediatric psychology. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 3-18). New York: Guilford.
- Thompson, R. J., Gustafson, K. E., Hamlett, K. W., & Spock, A. (1992). Stress, coping, and family functioning in the psychological adjustment of mothers of children and adolescents with cystic fibrosis. *Journal of Pediatric Psychology*, 17, 573-585.
- Varni, J. W., Blount, R. L., Waldron, S. A., & Smith, A. (1995). Management of pain and distress. In M. C. Roberts (Ed.), *Handbook of pediatric psychology* (2nd ed., pp. 105-123). New York: Guilford.
- Visintainer, M. A., & Wolfer, J. A. (1975). Psychological preparation for surgical pediatric patients: The effect on children's and parents' stress responses and adjustment. *Pediatrics*, 56, 187-202.
- Wallander, J. L., Varni, J. W., Babani, L., Banis, H. T., & Wilcox, K. T. (1989). Family resources as resistance factors for psychological maladjustment in chronically ill and handicapped children. *Journal of Pediatric Psychology*, 14, 157-173.
- Wolfe, V. V., Blount, R. L., Saylor, C. F., Dufour, D., & Finch, A. J. (1987). Nurse's perceptions of stress-reduction techniques in pediatric oncology. *Journal of the Association of Pediatric Oncology Nurses*, 4, 9-13.

24

Integrating Psychosocial Research and Practice in a Pediatric Hospital

ANNE E. KAZAK and ANNA T. MEADOWS

This chapter describes how psychosocial research and practice can be integrated in a pediatric hospital setting. We begin with background on the integration of psychology and pediatrics generally and then provide an overview of the social ecological approach that guides our work. The conceptual background, development and implementation of services, and changes over 11 years (1986–1997) of the Psychosocial Services Program in the Division of Oncology at the Children's Hospital of Philadelphia are then described in detail. The Psychological Services Program is a multidisciplinary program providing comprehensive mental health services for children with cancer and their families and conducting research related to the adjustment of children with cancer and their families. While our comments are focused on pediatric cancer within an academic setting, we have highlighted ways in which our efforts may be generalized to other clinical settings. In particular, we focus on the advantages of integrating research with clinical practice and the problems specific to medical settings, as well as proposed solutions.

INTEGRATION IN PEDIATRICS AND PSYCHOLOGY

Psychological research and practice has contributed to achievements in the broad areas of health and illness. Indeed, psychologists are well poised to bring an integrative focus of research and practice to all of their work and to contribute

ANNE E. KAZAK and ANNA T. MEADOWS • Division of Oncology, the Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104-4939.

collaborative expertise in medical settings with respect to clinical care, research, teaching and administration. However, there are many complex factors that affect how psychological practice has been integrated in pediatric settings. These include personal and setting characteristics, internal and external funding support for a sustained and directed program of research, established feasibility and efficacy of psychological interventions used, availability of psychologists trained for integrative pediatric psychology practice, and economic pressures related to changes in our national health care delivery systems and local responses to these forces. The challenge for pediatric psychologists is to maintain balance and respond creatively in the development, administration, and evaluation of services in the face of strong, persistent, fluctuating, and potentially disruptive forces.

While historically there has been a natural tendency for both psychologists and physicians to pursue either a "research" or a "clinical" career, these decisions increasingly have become complicated by economic considerations and availability of fewer opportunities for integrative careers. Aside from these larger issues, recent data suggest that integration may be enhanced by more collaboration and show that practitioners are more likely to read and utilize research than researchers are to actively solicit clinically relevant ideas for their research (Beutler, Williams, Wakefield & Entwistle, 1995). The Boulder model also remains the organizing framework for the majority of doctoral programs (O'Sullivan & Quevillon, 1992), suggesting that many new PhDs are well prepared for a scientist-practitioner career. Thus, although not necessarily a smooth process, psychologists have many of the basic skills necessary for an integrative approach.

The pursuit of one "track" or the other has implications for the advancement of psychology in medical environments. That is, clinically oriented psychologists are important members of treatment teams and provide effective assessments, interventions, and consultations based on psychological theory and research. They also may advance the training of clinical, health, and pediatric psychologists, and many pursue administrative responsibilities. Yet, the pressures to generate revenue for clinical services can limit psychologists' time and resources for innovative research or program development that is ultimately crucial to long-term established psychosocial care.

In contrast, research-oriented psychologists can pursue studies across a broad range of topics for which psychological data are important and gain the respect of medical colleagues as fellow scientists. They also can more readily achieve faculty appointments and advance academically. They tend to be, however, more removed from clinical services. This can be an obstacle in terms of translating research findings into practice, and with regard to educating staff as well as patients and families regarding the potential and actualized contributions of psychology in medicine. Of course, the difficulties of obtaining and maintaining external funding remain challenges throughout, and are the most pressing concern of researchers, across disciplines.

We maintain that the integration of research and clinical activities remains the most feasible strategy for advancing psychosocial care within medical environments. As practitioner-scientists in an academic environment psychologists, like their physician colleagues, conduct research and provide clinical services. In both academic and nonacademic settings, psychologists can offer unique research expertise that can advance clinical care or contribute creatively to evaluating ser-

vices. In either, the active clinical involvement of psychologists with other colleagues is likely to advance the development of clinical and research activities that are derived from important and basic questions of clinical care (e.g., pain, unusual patterns of neuropsychological test data, compliance, efficacy of interventions).

Pediatrics is a medical speciality that is welcoming and respectful of partnerships with psychologists. There is a large amount of overlapping interest in the treatment of and advocacy for children and families. Examples include the strong developmental training of both pediatricians and pediatric psychologists and an orientation towards competence rather than psychopathology. This allows for collaboration on many developmental, learning, and behavioral concerns of children and families and provides a natural fit with interventions that aim to promote the child and family's return to normal developmental pathways at times of acute or chronic illness. Both pediatricians and pediatric psychologists also recognize that diseases and treatments are not all like. One size does not fit all, particularly in understanding the impact of diverse illnesses at different stages of child and family development (Rolland, 1994). Similarly, both disciplines have long histories in advocacy for children and development of policies that promote children's health and well-being, and they support families in their caregiving responsibilities.

The fact that (the vast majority of) children live in families is an organizing framework for pediatricians and for psychologists with a family/systems orientation. Our programs are family-centered and based on a growing body of literature supporting the efficacy of family research in pediatric psychology (cf. Kazak, Segal-Andrews & Johnson, 1995c). Our research has shown that, contrary to previous "lore" which indicated that having a seriously ill child will erode family relationships, families generally cope competently with illness and treatment-related stressors, even if these adaptive solutions appear different from families of nonill children (cf., Kazak & Marvin, 1984; Kazak, 1989; Kazak & Simms, 1996). Yet, even with adaptive coping styles, without question serious pediatric illness is distressing for patients and families, and instances of maladaptive coping are not uncommon.

Indeed, children with serious health problems are at increased risk for mental health problems (Cadman, Boyle, Szatmar, & Offord, 1987; Lavigne & Faier-Routman, 1993). This well-established finding supports the need for psychosocial care for children in pediatric treatment facilities. Many clinical services are necessary and are provided routinely by pediatric psychologists (e.g., assessment, neuropsychological testing, consultation, psychotherapy). Psychologists also can provide the intervention and research expertise necessary to directly evaluate the impact of specific interventions and programs, building on the established efficacy of child and family interventions (Campbell & Patterson, 1995; Hibbs & Jensen, 1996).

Related to intervention is prevention: primary, secondary, and tertiary. Much of our work in pediatric oncology is geared toward preventing long-term adverse psychological effects; other areas of pediatrics also offer direct opportunities for involvement in primary prevention, such as human immunodeficiency virus-acquired immunodeficiency syndrome (HIV/AIDS), injuries, and adolescent pregnancy. A growing database of psychological research provides knowledge of family characteristics that are consistently associated with more adaptive outcomes (e.g., frequent and varied interactions, established community ties, encouragement of

autonomy and independence, creative problem solving, ability to adjust to change). Research also supports the importance of anxiety and patient and family subjective perceptions of the illness and treatment in predicting adjustment.

SOCIAL ECOLOGY

The social ecology of development, proposed by Bronfenbrenner (1979), provides the developmental foundation for a model that we have applied to serious childhood illness with regard to research, intervention, and the conceptualization of mental health services for children and families. It is a framework for understanding the interactions among childhood illness and the individuals and systems internal and external to the family. Social ecological theory also establishes the importance of reciprocity and change in the course of development. As such, it offers a framework that can help organize our understanding of the ways in which children, families, illnesses, and other systems interact. It also provides an organization that can guide family systems interventions that promote competence and positive developmental outcomes.

Social ecology examines relationships between the developing individual and the settings and contexts in which the person functions actively (Bronfenbrenner, 1979). For children with serious illness these settings include homes, schools, neighborhoods, communities, and multiple facets of our larger society. Each context in itself is rich and complex and each has been studied extensively. Social ecological theory moves beyond looking at individual settings by positing that the interconnections among settings impact on child development. Also basic to social ecological theory is that events that occur in settings in which the child is not present can have significant impact on the child's development.

In social ecology, the child is at the center of nested concentric spheres of influence. More detailed descriptions of the model can be found elsewhere (Bronfenbrenner, 1979; Kazak et al., 1995b). In Table 1 we summarize some of the most pertinent principles of social ecology, focusing on those that are most directly related to our model of clinical services. Briefly, the family is the system most immediate to the child. The presence of parents, siblings, and extended family is integral to the child's ongoing care and development. It is the cornerstone of our

Table 1. Principles of Social Ecology

The child or adolescent with pediatric illness is affected by multiple, complex aspects of his or her context, which can be understood as a series of concentric circles, with the child at the center.
The family is the primary system. Other crucial systems include schools, peer groups, and neighborhoods.
A developmental approach encourages a view of children and families as competent.
The process of assessment and intervention should include multiple persons and settings.
Health care settings and the individuals in them are key in promoting growth and development.
Systems that have indirect effects on children include parental workplaces, parents' social networks, and the contexts in which siblings function.
Broader systems (e.g., laws, policies, cultural, subculture) are important.
Development rests on interpersonal relationships and opportunities provided within settings.

clinical and research approach. We also consider the disease to be part of the most immediate system. Diseases (and their treatments) place particular and specific demands on the child and family and at certain times can become the fulcrum of their own patterns of activity, often organizing family members' behavior.

Social ecology examines the interrelationships among systems that impact on children, for example, interactions among families, schools, hospital, and communities. While these interrelationships are important for all children and adolescents, families of ill children often have short- and long-term educational needs that accentuate the importance of family–school interfaces. Moreover, there are short- and long-term relationships with the members of the health care team that change, with the nature and demands of treatment. The teams and settings themselves also change, and must change as the child grows into later adolescence and young adulthood. Perhaps because of the inherent complexity of these interrelationships, relatively little is known empirically about how the child navigates among systems and how these changes and transitions impact adjustment. The concept of a “system of systems” provides a helpful framework for conceptualizing clinical intervention.

Moving into the realm of indirect effects on child development, environments that are important to the child’s development include parental workplaces, parental social networks, and the classes and peer groups of siblings. Understanding and intervening to strengthen existing relationships and build new partnerships with families of seriously ill children is an example of social ecological practice. The model provides a way in which we can think about the influence of larger systems, such as subcultures and culture, in our work.

Social ecology emphasizes the importance of individual development and the implications of developmental processes for coping and adapting. As such, attention is directed toward how children grow and expand, differentiating in their evolving abilities and competencies in mastering new environments and changes over time. Illness and treatment can have many undeniable effects on all realms of growth and development. However, maintaining an adaptive, competency-based framework helps promote “normal” development in the face of adversities associated with illness and treatment. The individuals and systems that are inherent throughout the social ecological context provide the key interpersonal relationships that promote this development.

Change is an integral part of the growth of all systems. Social ecology provides a normalizing perspective on change and transitions that is helpful in understanding how children, families, and health care systems interact over time. When a child becomes ill or is diagnosed with a serious illness, the disease, treatment, and necessary emersion in medical care systems tend to dominate and reorganize child and family functioning. That is, while serious childhood illness can prompt some of the most striking family transitions and is clearly abnormal (in the sense that children are not “supposed to” have life-threatening illnesses or need to endure illness-related struggles), transitions themselves are among the most natural of experiences for children and adults. Viewing serious pediatric illness as a transition allows for a normalization of these inherently complex and often unexplainable events. For example, the ways in which families have coped with prior unexpected circumstances and/or their general level of comfort with the emotions and processes associated with change and transition (loss of control, uncertainty,

challenge, sadness, fear, hope) can be beneficial in coaching children and families throughout the course of treatment.

As is supported by research (of our team and others), families of ill children are "normal" families confronting stressors that challenge their usual patterns of behavior and coping styles. Despite these circumstances and their potential for short- and long-term distress, most families cope adaptively and most children are able to grow up with well-defined strengths and abilities to face subsequent hurdles of adult life. The process of development is enhanced by interactions with persons and environments that promote trust, competence, respect, and support. For some children and families, this process may necessitate actively averting or correcting developmental derailments. For others, it may provide the opportunity to refine or reframe their general orientation to incorporate the illness and treatment as part of themselves and their families. The clinical and research programs that we describe are built on these aspects of social ecology.

THE PSYCHOSOCIAL SERVICES PROGRAM

In this section we describe briefly the Division of Oncology at the Children's Hospital of Philadelphia (CHOP) and the history of psychosocial care, the transition to the current model, and the organization of the program, including clinical services and research programs.

The Division of Oncology at CHOP is a large children's cancer treatment center, with approximately 300 new patients seen each year. With increasing rates of survival for childhood cancer, the number of patients followed regularly, both on and off treatment, is also increased and now numbers more than 2500. Attending oncologists are faculty members in the Department of Pediatrics at the University of Pennsylvania, as is the Director of Psychosocial Services. CHOP is an independent hospital and other nonfaculty staff (including the psychosocial services staff) are employed by the hospital.

As with most medical settings, the systematic incorporation of psychology (excluding the role of psychological evaluation) is relatively recent. Beginning in the 1970s, psychosocial care was provided by social workers. By the mid-1980s, the social work staff numbered five. The emphasis was on clinical social work, with psychiatric consultation-liaison through Philadelphia Child Guidance Center, at that time an affiliated hospital. These services provided a strong basis for intensive psychosocial services for children with cancer. However, psychosocial research was absent and a sense developed that clinical services could be expanded using a different model.

In 1986, together with a strong emphasis on long-term survivors of childhood cancer and initial funding from a research grant to one of the authors (ATM), a psychologist (AEK) was recruited to work full-time and primarily devoted to research within the Division of Oncology. A PhD special educator, also grant-funded, joined the team to develop research and clinical interventions related to the long-term impact of cranial irradiation and chemotherapy on learning. With a change in medical leadership in the Division of Oncology in 1989 (ATM), the Psychosocial Services Program was established in order to integrate clinical services and research in mental health. The psychologist was appointed to direct this program. The sustained emphasis of our psychosocial programs on cancer sur-

vivors reflected the work initiated and subsequently developed by the medical leadership (ATM).

The transition in the philosophy and organization of psychosocial care resulted in a difficult 2-year period. As a consequence of the changes we describe, all the social work staff resigned. In order to avert a negative impact on clinical care, our emphasis during this period was necessarily twofold: to recruit staff to develop the emerging program and to provide the clinical services necessitated by the resignations. These interim activities were critical background steps in order to establish the later research orientation. The relationship between social work and psychology was instructive here, in part because it underscores the dangers of dichotomizing research and clinical services and the challenges of multidisciplinary work. Prior to her appointment as director of the program, the psychologist was seen as a "researcher," and therefore had a circumscribed role in clinical care, the majority of which was provided by the social workers. While relationships were cordial and collaborative on patient care, the underlying tension between research and clinical services came to the forefront with the change in leadership and introduction of a new agenda.

Structure of the Psychosocial Services Program

Although the size of the program has fluctuated with changes in grant funding, during the period under consideration, the Psychosocial Services Program (PSP) consisted of 12 full time staff members: four PhD psychologists, six master's-level social workers, and two child life specialists. Trainees across the disciplines also have been active members of the program staff. We were pleased that the program received recognition and resulting dissemination, both within CHOP (Bonnem, 1995) and in the pediatric psychology community (Drotar, 1995).

The structure of the program incorporates a matrix supervisory-management system. PSP members are responsible to both the director of the PSP and to their department head within the hospital. The director of the PSP and the involved department heads (social work, child life) have developed policies to guide the supervision of staff. This structure allows staff to have a close association with a multidisciplinary team in the Division of Oncology while maintaining their affiliations with their specific discipline within the hospital. As there is no psychology department in the hospital, psychologists report only to the director of the PSP.

With respect to funding, social work positions have been supported, as in most hospitals, mainly by the Department of Social Work. Donations to the Division of Oncology have been instrumental in developing and maintaining child life programs. Psychologist salaries have been derived from grants, divisional funds, and clinical services. The advantages of external funding include development of an academically oriented program in which staff have time protected to conduct research and the autonomy that accompanies external funding. The most evident downside is the fluctuations in obtaining such resources, with attendant increases and decreases in staff and programs.

Clinical Services

Clinical care is the top priority for the PSP. In Table 2 we summarize the basic clinical services provided, with brief descriptions of each. They are quite similar to

Table 2. Examples of Clinical Services: The Psychosocial Services Program in the Division of Oncology at The Children's Hospital of Philadelphia

Primary psychosocial care	Every newly diagnosed child with cancer is assigned a member of the psychosocial staff who joins the treatment team. This social worker or psychologist conducts an initial psychosocial evaluation of the child and family and provides support during the initial diagnostic workup and communication of diagnosis and treatment plan to patient and family. Early psychosocial assessment of the child and family allows for determination of competencies and risk factors. The staff member collaborates with the family over the course of treatment, both inpatient and outpatient.
Resources and community linkages	As part of the family assessment process, the need for access to community resources is evaluated. We provide referrals to local and national organizations as needed and also have prepared a handbook for parents. The school reintegration program provides for school visits from nursing and/or psychosocial staff for newly diagnosed children.
Psychological consultation	Referrals to psychologists are made for inpatient and outpatient behavioral, emotional, or family problems.
Support and parent education	Parent groups are held weekly on the inpatient unit. Monthly support groups are held for some specific diseases and parent programs in the outpatient clinic are also held.
Parent to Parent Support Network	Peer support is provided. Support parents complete an 18-hour training program.
Sibling program	Programs for siblings are offered during which siblings learn about cancer and have the opportunity to discuss their concerns, engage in medical play, and interact with their parents in programs that are family oriented.
Child Life Program	Inpatient and outpatient programs offer therapeutic and diversionary activities.
Neuropsychological testing and psychoeducational consultation	Comprehensive neuropsychological evaluations are offered prior and subsequent to bone marrow transplantation and total body irradiation and cranial irradiation for patients on select treatment protocols and for evaluation of learning difficulties that may be related to treatment.
Communication with other disciplines	Weekly multidisciplinary rounds are held. Psychosocial staff are assigned to particular treatment teams (e.g., BMT, neurooncology). A newsletter, the Psychosocial Services Times (PSST) is distributed to staff and families and provides updated information on psychosocial programs and issues.
Psychotherapy	Individual and family therapy is provided, as needed.

currently recommended standards for psychosocial care in pediatric oncology, with the goals of "minimizing emotional stress and its associated psychosocial morbidity" (Noll & Kazak, 1997). As we will discuss later in the chapter, current changes in health care have impacted the services we provide. These changes challenge us to find ways in which we can work collaboratively to restructure some of our approaches. The goal is to maintain a high level of care, keeping the standards prominent but realizing that new creative models may prove to be as efficacious, even more cost-effective, collaborative, and consistent with family-centered care models in pediatrics. For example, the primary psychosocial care approach (Table 2) is a comprehensive, preventive one that has been well received by patients, parents, and staff alike. Our approach challenged some of the traditional professional boundaries between psychologists and social workers. Families

may be assigned to either. This spreads the workload and helps to maintain a sense of teamwork and camaraderie within the psychosocial team. The expectation is that the expertise needed is within the team (rather than in each individual) and that team members turn to others for more specialized assistance when needed. This model reflects the belief that childhood cancer is an extreme stressor. The disease and treatment-related problems impact on children and families and can best be addressed within a treatment context including individual, family, community, and systems issues.

However, the “blurring of boundaries” initially baffled some staff members who felt uncomfortable with “psychologists doing social work” or were reluctant to believe that the social worker could successfully provide more intensive levels of intervention, even with a psychologist as “backup” consultant. The model has prompted social workers to become more involved in research and has encouraged them to play a more central role in clinically complex cases. It also sometimes has been difficult for psychologists to accept the model, particularly those whose training did not emphasize interdisciplinary work and whose programs of research were still in early stages of formulation. Supervisors of graduate student, intern and postdoctoral psychologists, understand the need to advocate for development of a breadth of psychosocial skills and to refine and solidify research interests as important aspects of the mentoring process.

Underscoring the multidisciplinary nature of the PSP, we developed the first outpatient child life program at CHOP, expanding on the well-established inpatient program. The director of the PSP is the direct supervisor for the outpatient child life specialist and, along with a multidisciplinary team in the Division of Oncology, developed the program. This has allowed for integration of the therapeutic and diversionary activities of child life with other psychosocial members’ treatment plans and has provided opportunities for child life specialists to receive additional training, as they interact with other members of the PSP team.

It is beyond the scope of this chapter to describe all the clinical services in detail. The reader will find brief descriptions in Table 2 or may contact the authors for more information.

Research

Conducting research that will advance understanding of how children and families experience illness and treatment and the types of intervention that may promote successful adaptation is a critical component of our program (Table 3). Our work is guided by a family/systems framework, consistent with a social ecological perspective (Kazak, 1989). The importance of evaluating what we do and helping to promote knowledge in our field is unquestioned in our academic setting, for members of all disciplines. In the following discussions, we present both the content of our major research efforts and reflections on the process.

As with all complex systems, many factors influenced the course of the program’s development. Some decisions were based on a purposeful choice (e.g., that we would emphasize cancer survival). Others were more opportunistic (e.g., a call for proposals in an area in which we felt able to respond). Many situations were out of our control (e.g., one proposal funded, another not). There were more “dead ends” encountered than could reasonably be described in a brief chapter. Across

**Table 3. Examples of Research:
The Psychosocial Services Program in the Division of Oncology at The Children's Hospital of Philadelphia**

Area of research	Funding/staffing	Dissemination	Clinical links
Psychological sequelae of childhood cancer survival	National Cancer Institute University of Pennsylvania Cancer Center and School of Nursing Psychology, psychiatry, pediatric oncology, nursing	Kazak et al. (1997) Kazak et al. (1996d) Stuber et al. (1996) Kazak & Barakat (1997) Stuber et al. (1997)	Parents/survivors continue to experience symptoms of posttraumatic stress after treatment ends Risk factors can be identified for survivors and their families Advancement in the assessment of posttraumatic stress related to pediatric illness Currently developing interventions to be evaluated
Pain and distress during medical procedures	National Cancer Institute Psychology, pediatric oncology, nursing, anesthesiology, social work, child life	Kazak et al. (1995b) Kazak et al. (1995a) Kazak et al. (1996a) Kazak et al. (1996c) Kazak et al. (1996b)	Combined pharmacological-psychological approaches to pain/distress during are effective Advancement in the assessment of procedural distress for parents and staff There are fluctuations in the acceptance of new programs which can be studied quantitatively
Neuropsychological outcome in bone marrow transplantation		Simms et al. (1996)	Why clinical recommendations are not more often translated into practice is worthy of further study Neuropsychology evaluations are important in understanding late effects of BMT
Neurooncology research	Psychology, pediatric oncology, neurology, social work, epidemiology	Bunin et al. (1995) Radcliffe et al. (1996) Foley et al. (in press)	Feedback process for study participants can be refined Competencies can be identified in many survivors of pediatric brain tumors and their families In some brain tumor patients, risk is higher and associated with family functioning

time, however, establishing and maintaining collaborative relationships within the hospital, university, and broader national and international communities proved fruitful.

An integrated clinical research program requires attention to administrative and organizational concerns. That is, funding for research is a necessary prerequisite for a successful program; however, there are many additional and sometimes more mundane issues that are important. For example, actively interacting with larger organizational structures (in this case, the Division of Oncology, CHOP, and the University of Pennsylvania) positions a program to benefit from resources related to seeking funding, secretarial support, collaboration with others, and trainees. As a further illustration, although we do not maintain a separate PSP database for research, we routinely use the resources of the tumor registry in the Division of Oncology, which is supported in part through the national interinstitutional studies of the Children's Cancer Group (CCG).

Ideas for projects and momentum related to them can come from many different persons and settings within the PSP and broader systems. Some of these mechanisms are illustrated below in the discussion of specific projects. Reflecting our pragmatic approach, most projects and questions evolved directly from clinical concerns. Ongoing staff involvement was similarly sustained by a sharp clinical focus. For example, a social worker currently actively involved in the development of an intervention for cancer survivors had expressed an interest over time in learning more about intervention packages and how they could be applied to our patients.

As noted above, the intent to expand psychosocial research was central in the decision to incorporate psychology in our setting. Already found in previous research at CHOP was the association of cranial irradiation with later declines in intellectual performance (Meadows et al., 1981; Peckham, Meadows, Bartel, & Marrero, 1988). This work, initially funded by a core grant from the National Cancer Institute and later by a private foundation, was one of the pioneering reports suggesting the importance of expanding psychosocial care and research to children who have survived cancer and their families. During her first few years at CHOP, one of the responsibilities of the psychologist (AEK) was to complete the studies begun earlier. That is, data on neuropsychological functioning and psychological adjustment had been collected and needed to be analyzed and disseminated.

Our subsequent work (described in more detail below) reflects the shared beliefs that, as the number of childhood cancer survivors would increase, the importance of determining psychosocial sequelae would also increase. The continued consideration given to childhood cancer survivors also illustrates the process of interdisciplinary collaboration. That is, common interests (research questions, clinical concerns) approached from diverse angles over time allowed us to expand overall clinical and research in this area. We also maintained a consistent orientation toward practical research questions and sought to translate clinically important issues into research projects and to translate the findings of research into our clinical care.

Another key collaboration evolved with funding from the US Bureau of Maternal and Child Health. The principal investigator (Beverly Lange, MD) worked with a multidisciplinary team to establish the safety of chemotherapy given at home and also supported the importance of educational intervention for children with cancer

who had received cranial irradiation (Close, Burkey, Kazak, Danz & Lange, 1995). In submitting a new grant to the funding agency, we built on this work and bridged to a more psychosocial framework, within a family-centered, community-based model. This grant funded clinical projects, including the Parent to Parent Support Network (Table 2) and a similar program for peer support among patients, and provided for expansion of both clinical service and research in the area of neuropsychology. It also allowed for expansion of an annual workshop program for teachers of our patients and our patients' siblings at the hospital, thereby strengthening our linkages with schools and communities.

Long-term Effects

As noted earlier, a major research effort continues to be understanding the psychological sequelae of childhood cancer survival for children and families (Table 3). While the potential adverse effects of cancer and its treatment on cognitive and learning ability were widely accepted from previous research, the long-term psychological impact these experiences may have for patients and their parents was less well understood. We conducted two studies that examined differences between childhood cancer survivors and comparison group families (Greenberg, Kazak, & Meadows, 1989; Kazak & Meadows, 1989). The data were consistent with previous work in highlighting the generally adaptive functioning of families (Kazak, 1989). However, subsequent studies support endorsement of more distressing symptoms in mothers and fathers of childhood cancer survivors than in parents of never-ill comparison group children. While there is evidence for similar symptoms in survivors themselves, the data are less clear than for the parents (Kazak & Barakat, 1997; Kazak, et al., 1997; Kazak, Stuber, Barakat, & Meeske, 1996d; Stuber et al., 1997).

These findings have direct clinical implications that we have utilized. For example, our findings about posttraumatic stress are used in discussions with families nearing the end of treatment. We have similarly expanded our work to include the development of an interview methodology to clarify potentially traumatic aspects of cancer and its treatment so that we can develop ways of preventing long-term distress. Our current efforts in this area are directed toward designing and evaluating an intervention that combines cognitive-behavioral and family approaches.

Procedural Pain

Another important area of research is that of pain and distress during invasive medical procedures (e.g., lumbar punctures and bone marrow aspirates and biopsies) in the treatment of leukemia. This is a well-established area of research for pediatric psychologists and we discuss below our project in this area. It also serves to illustrate the process of developing such a study within a medical unit and the emphasis on staff (as well as patient and family outcome) and concern for the broader issues of mechanisms of change in pediatric practice.

The project emanated from a multidisciplinary committee (Clinical Care Committee) and was driven by concerns that "we could do better" in handling children's pain and distress during invasive medical procedures. Psychological input

was important in designing psychological interventions (e.g., using distraction, relaxation, hypnosis, etc.). However, clarifying the ways in which research could be conducted to answer relevant questions was probably the most crucial step in terms of establishing a firm multidisciplinary approach to this topic and was critical in obtaining funding. We conducted a prospective study comparing a pharmacological intervention with a combined family focused pharmacological-psychological approach. Called the Analgesia Protocol for Procedures in Oncology (APPO), the program also addressed a unique systemic question, in addition to the usual behaviorally oriented outcomes in this field of research. That question was: *What is the effect of introducing change in the form of an organizationwide pain protocol on patients, parents, and staff?* In order to address this question, the prospective data collection included staff as well as families and addressed issues related to the use and acceptance of a new approach to clinical care as well as the specific outcomes for children and families.

We identified four tenets underlying our approach to this program that can be generalized as important components of our overall program: (1) families are the most important force in helping children cope with serious illness; (2) complex medical problems can be effectively addressed through integrated, multidisciplinary teamwork; (3) change demands a systemic focus and is characterized by alternating periods of achievement and resistance; and (4) research on the process and outcome helps foster acceptance of change. Overall, the program helped establish the steps that are important to successfully introducing change in clinical practice. Details may be found in the series of project papers cited in Table 3.

Similar to the discussion of clinical care, full discussion of all of our research activities is beyond the scope of the current chapter. Table 3 summarizes some of the major research activities of our team.

Education and Training

Our PSP education and training activities include physicians and multidisciplinary mental health professionals. Postdoctoral fellows, graduate students from local clinical, counseling, and school psychology programs, and interns electing oncology rotations from predoctoral internship programs receive psychological training in our division. Their training encompasses all PSP activities and they experience the blend of research and clinical pursuits seen throughout the program. Social work and child life interns from local and national programs also complete rotations and placements in the PSP.

Educational and training activities that interface with physicians include weekly psychosocial rounds with the inpatient multidisciplinary team, participation in journal clubs, and monthly psychologist-led groups for faculty discussions of teamwork and the ways in which their work affects their lives and relationships. This program also has been implemented for hematology-oncology fellows.

Research activities reflecting a mentor relationship also have been successful in the PSP. For example, a medical student completed a full-time research fellowship during a “year-out” program and a social worker conducted a research study with the guidance of a psychologist. Finally, we also have published recommendations for teaching and translating the social ecological model in clinical treatment (Kazak & Simms, 1996; Simms, 1995; Simms & Kazak, 1998).

MAINTAINING AN INTEGRATED PROGRAM OF RESEARCH AND PRACTICE

In this concluding section we provide summaries of some of the characteristics of our programs and processes over time that have been helpful in sustaining an integrated program of research and practice.

Allow Time

Time is necessary to establish viable programmatic research. Effective collaborations and support must be obtained from leadership in order to assure investment in the program. Pilot studies are crucial and may have to be done with minimal staff and resources. Educating others about psychological research and engaging their active participation are also important. For example, the length of time needed for a complete neuropsychological battery is much longer than that required for many medical tests. Physicians and nurses discussing testing with families must take an educational approach, assuring that the time required for the procedure is clear to the family, as well as the importance of conducting the evaluation.

Change Is Inevitable

Research funding inevitably waxes and wanes and its cycles involve lengthy processes. A critical component of effective program administration is to foresee points at which existing sources of funds may be reduced (e.g., grants end, federal and foundation funding priorities change). Recently, changes of this type have occurred concurrently with alterations in hospital policies that can decrease staff (e.g., staff who resigned or moved from full-time to part-time were not replaced), with uncertainty promoted by the loss of funding sources (e.g., efforts to obtain funding at the state level were affected by changes in political appointments and general economic uncertainty increased demand on limited private and foundation monies), and with changes associated with managed care. While each health care facility and system has its own circumstances that affect the structure and stability of psychosocial programs, we have outlined some of the strategies that we have found helpful in maintaining our programs during periods of uncertainty.

Pursue More External Funding while Seeking Other Resources

Perseverance with regard to grant submissions is critical. This includes resubmissions of grants with changes requested by review committees, identification of new sources of public and private support, and exploration of interim sources of funding from internal mechanisms. These processes are labor and time intensive, exhausting, and not necessarily logical or productive. Advocating for psychological research, particularly as applied to questions that are not primarily psychological should be translated into support for psychologists on the grants of medical and nursing colleagues.

Strengthen Collaborative Links with Hospital Administration

We saw the opportunity for closer collaboration with the leadership within the hospital and ways in which the Director of Psychosocial Services could have more direct association with hospital departments and leadership. This helped articulate a blueprint for future changes that could include more flexibility in procurement and distribution of funds. That is, after initial discussions with vice presidents who oversaw relevant departments, a workgroup was formed, including the Director of Psychosocial Services and leaders from social work, nursing, and child life. Our common interests and mission were clear, as was the respect for the program that had been developed within the Division of Oncology. We were able to develop a plan that focused on a few key questions (e.g., What services are necessary? What are the outcomes of psychosocial interventions? How can we distribute services over the continuum of care?). Finally, we obtained explicit support and agreement at the hospital level for a psychosocial intervention approach that incorporated a developmental, contextual, social ecological model.

Frame Clinical Services within a Clear Model and Prioritize

The social ecological framework (Kazak, 1989) that guided our research is also relevant to clinical practice. That is, the contextual approach, combined with a strong developmental framework, is central to our clinical approaches. Applying this model more systematically helps organize psychosocial, medical, and nursing staff to understand services provided to patients, parents, siblings, extended families, and schools with a contextual model and to attend closely to developmental concerns. While the level of care that we have provided is relatively intensive and complex, several questions were posed that helped in terms of prioritizing services, particularly during periods in which fewer resources were available. These include: (1) What services are necessary for patients and families and which are recommended but optional? (2) What are the outcomes of our psychosocial services, for patients, families, and staff? (3) As patient and family reactions to illness and treatment change over time, how can our services vary to accommodate periods in which receptivity is greatest? and (4) How can our roles as psychosocial staff be maintained and expanded with a collaborative consultant model so that the clinical questions raised by staff as well as families are addressed and resolved?

Reaffirm Multidisciplinary Workgroups on Research and Clinical Service

During times of retrenchment there is a natural tendency for individuals or groups defined by disciplines to become more isolated in their work. That is, with fewer resources, people tend to focus on those aspects of their role that they feel are most fundamental to their discipline. In the PSP we saw this process in terms of psychologists spending relatively more time writing grants and completing research projects and social workers focusing on the daily demands of clinical care.

However, basic to the notions of multidisciplinary teamwork and progress toward our long-term goals, we reaffirmed our belief in active multidisciplinary projects and facilitated projects that would promote our long-term goals. Two examples are used to illustrate this. First, although federal funding for psychologi-

cal intervention outcome research for survivors and their parents was pending, divisional funding was provided to conduct a pilot for the project. Staff from multiple disciplines (psychologists, oncologists, a nurse practitioner, a social worker, and trainees from these disciplines) met weekly, manualized the program, and conducted a pilot, the results of which were incorporated into the resubmission of the grant.

Second, providing psychosocial services to a steady number of patients with fewer staff had necessitated some streamlining of care. A team of psychosocial staff (including, in addition to psychologists, social workers, and child life specialists, a nurse and a hospital teacher) initiated the development of a systematic assessment tool. That is, while there was a high level of consensus with regard to what information is important in helping families, our assessment processes were diverse and noncoordinated. The new assessment tool—process can guide and evaluate clinical services (e.g., what types of psychosocial care should be provided) as well as answer basic research questions related to the process of adjustment to serious pediatric illness over time to serious pediatric illness (e.g., to what extent are specific psychosocial symptoms or social support at diagnosis related to longer-term adjustment). In this way, it provides a strong example of how clinical care can be facilitated and can generate translational data and can guide future clinical services.

SUMMARY

The Psychosocial Services Program in the Division of Oncology at the Children's Hospital of Philadelphia is a multidisciplinary program of clinical services and research. In this chapter we have described the background guiding the program and its history and structure and have provided examples of approaches that we have found helpful in sustaining the program over time. Our experiences have indicated that development, refinement, and sustenance of a program takes from 5 to 10 years, highlighting the importance of persistence, flexibility, and careful attention to broader systemic issues of medical settings. The five ingredients of successful programmatic collaboration outlined by Drotar (1995, pp. 104–107) are evident in our work. These include: departmental recognition of psychologists' contributions, clarity of administrative communication and decision making, leadership and power, teamwork and group support, and developing economic resources. The Psychosocial Service Program also maintains an orientation of integrating research and clinical intervention that has necessitated the establishment of mechanisms that can sustain research activity during periods of natural fluctuation in external sources of funding. Maintaining a clear focus on the needs of patients, families, and staff provides direction for clinical services and helps refine research initiatives that will promote the integration of psychosocial and broader pediatric care.

REFERENCES

- Beutler, L., Williams, R., Wakefield, P., & Entwistle, S. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist*, 50, 984–994.
Bonnem, S. (1995). *Two decades: The Children's Hospital of Philadelphia, 1974–1994*. Philadelphia, PA: Children's Hospital of Philadelphia.

- Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Bunin, G., Kazak, A., & Mitelman, O. (1995). Informing subjects of epidemiologic study results. *Pediatrics*, 97, 486-491.
- Cadman, D., Boyle, M., Szatmari, P., & Offord, D. (1987). Chronic illness, disability, and mental and social well-being: Findings of the Ontario Child Health Study. *Pediatrics*, 79, 505-512.
- Campbell, T., & Patterson, J. (1995). The effectiveness of family interventions in the treatment of physical illness. *Journal of Marital and Family Therapy*, 21, 545-583.
- Close, P., Burkey, E., Kazak, A., Danz, P., & Lange, B. (1995). A prospective controlled evaluation of a homecare program for chemotherapy and antibiotics in children with cancer. *Pediatrics*, 95, 896-900.
- Drotar, D. (1995). *Consulting with pediatricians: Psychological perspectives*. New York: Plenum Press.
- Foley, B., Barakat, L., & Herman-Liu, A. (in press). The impact of childhood hypothalamic/chiasmatic brain tumors on child adjustment and family functioning. *Children's Health Care*.
- Greenberg, H., Kazak, A., & Meadows, A. (1989). Psychological functioning in 8-16 year old survivors of childhood cancer. *Journal of Pediatrics*, 114, 488-493.
- Hibbs, E., & Jensen, P. (1996). *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice*. Washington, DC: American Psychological Association.
- Kazak, A. (1989). Families of chronically ill children: A systems and social ecological model of adaptation and challenge. *Journal of Consulting and Clinical Psychology*, 57, 25-30.
- Kazak, A., & Barakat, L. (1997). Brief report. Parenting stress and quality of life during treatment for childhood leukemia predicts child and parent adjustment after treatment ends. *Journal of Pediatric Psychology*, 22, 749-758.
- Kazak, A., Barakat, L., Meeske, K., Christakis, D., Meadows, A., Casey, R., Penati, B., & Stuber, M. (1997). Posttraumatic stress, family functioning, and social support in survivors of childhood leukemia and their mothers and fathers. *Journal of Consulting and Clinical Psychology*, 65, 120-129.
- Kazak, A., Blackall, G., Boyer, B., Brophy, P., Buzaglo, J., & Himelstein, B. (1996a). Implementing a pediatric leukemia intervention for procedural pain: The impact on staff. *Families, Systems and Health*, 14, 43-56.
- Kazak, A., Boyer, B., Brophy, P., Johnson, K., Scher, C., Covelman, K., & Scott, S. (1995b). Parental perceptions of procedure-related distress and family adaptation in childhood leukemia. *Children's Health Care*, 24, 143-158.
- Kazak, A., Blackall, G., Himelstein, B., Brophy, P., & Daller, R. (1995a). Producing systemic change in pediatric practice: An intervention protocol for reducing distress during painful procedures. *Family Systems Medicine*, 13, 173-185.
- Kazak, A., & Marvin, R. (1984). Differences, difficulties and adaptation: Stress and social networks in families of handicapped children. *Family Relations*, 33, 67-77.
- Kazak, A., & Meadows, A. (1989). Families of young adolescents who have survived cancer: Social emotional adjustment, adaptability, and social support. *Journal of Pediatric Psychology*, 14, 175-191.
- Kazak, A., Penati, B., Boyer, B., Himelstein, B., Brophy, P., Waibel, M.K., Blackall, G., Daller, R., & Johnson, K. (1996b). A randomized controlled prospective study of a psychological and pharmacological intervention protocol for procedural distress in pediatric leukemia. *Journal of Pediatric Psychology*, 21, 615-631.
- Kazak, A., Penati, B., Waibel, M. K., & Blackall, G. (1996c). Perceptions of Procedures Questionnaire (PPQ): Psychometric properties of a brief parent report measure of procedural distress. *Journal of Pediatric Psychology*, 21, 195-207.
- Kazak, A., Segal-Andrews, A., & Johnson, K. (1995c). The role of families and other systems in pediatric psychology research and practice. In M. Roberts (Ed.). *Handbook of pediatric psychology* (2nd ed., pp. 84-104). New York: Wiley.
- Kazak, A., & Simms, S. (1996). Children with life threatening and special health care needs: Psychological difficulties and interpersonal relationships. In F. Kaslow (Ed.). *Handbook of relational diagnosis* (pp. 225-237). New York: Wiley.
- Kazak, A., Stuber, M., Barakat, L., & Meeske, K. (1996d). Assessing posttraumatic stress related to medical illness and treatment: The Impact of Traumatic Stressors Interview Schedule (ITSIS). *Families, Systems and Health*, 14, 365-380.
- Lavigne, J., & Faier-Routman, J. (1993). Correlates of psychological adjustment to pediatric physical disorders: A meta-analytic review and comparison with existing models. *Journal of Developmental and Behavioral Pediatrics*, 14, 117-123.
- Meadows, A., Gordon, J., Massari, D., Littman, P., Ferguson, J., & Moss, K. (1981). Declines in IQ scores

- and cognitive dysfunctions in children with acute lymphocytic leukemia treated with cranial irradiation. *Lancet*, 2, 1015–1018.
- Noll, R., & Kazak, A. (1997). Psychosocial care. In A. Ablin (Ed.), *Supportive care of children with cancer: Current therapy and guidelines from the Children's Cancer Group* (pp. 263–273). Baltimore: Johns Hopkins Press.
- O'Sullivan, J., & Quevillon, R. (1992). 40 years later: Is the Boulder model still alive? *American Psychologist*, 47, 67–70.
- Peckham, V., Meadows, A., Bartel, N., & Marrero, O. (1988). Educational late effects of childhood acute lymphocytic leukemia. *Pediatrics*, 81, 127–133.
- Radcliffe, J., Bennett, D., Kazak, A., & Phillips, P. (1996). Adjustment in childhood brain tumor survival: Child, mother, and teacher report. *Journal of Pediatric Psychology*, 21, 529–540.
- Rolland, J. (1994). *Families, illness and disability: An integrative treatment model*. New York: Basic Books.
- Simms, S. (1995). A protocol for seriously ill children with severe psychosocial symptoms: Avoiding potential disasters. *Family Systems Medicine*, 13, 245–257.
- Simms, S., & Kazak, A. (1998). Family systems interventions in pediatric neuropsychiatric disorders. In C. E. Coffey & R. Brumback (Ed.). *Textbook of pediatric neuropsychiatry* (pp. 1449–1464). Washington, DC: American Psychiatric Press.
- Simms, S., Kazak, A., Gannon, T., Goldwein, J., & Bunin, N. (1996). Assessing neuropsychological outcome in bone marrow transplantation. Under editorial review.
- Stuber, M., Christakis, D., Houskamp, B., & Kazak, A. (1996). Posttraumatic stress in survivors of childhood leukemia. *Psychosomatics*, 37, 254–261.
- Stuber, M., Kazak, A., Meeske, K., Barakat, L., Guthrie, D., Garnier, H., Pynoos, R., & Meadows, A. (1979). Predictors of posttraumatic stress in childhood cancer survivors. *Pediatrics*, 100, 958–964.

VII

Integrating Research and Policy

Another important but neglected area concerning the integration of research in pediatric and clinical child psychology concerns applications of research to policy. Research in pediatric and child clinical psychology generally has been isolated from policy considerations despite the fact that such research bears significant relevance to policy applications. In Chapter 25, Wells considers the ways that research can be used to enhance mental health practice and policy and considers obstacles to utilization of research in practice, for example, include research designs that cannot be easily generalized, concepts that are not relevant to practice situations, and data analytic approaches that do not consider clinical significance. Wells also considers the differences in the organizational demands and contexts in which various professionals function and makes several recommendations to enhance appropriate use of research by policy makers and training methods to enhance psychologists' ability to integrate research with practice. Wells highlights the need for researchers to learn how to focus effectively on the dissemination of research and to develop recommendations for improving policy and practice based on their research findings.

25

Use of Research in Mental Health Practice and Policy

KATHLEEN WELLS

INTRODUCTION TO THE ISSUE

The view that research can be used to improve the human condition is widely held in American society. This view is compatible with the empirically minded spirit in which the nation was founded (Lynn, 1978), with the American tendency to resist making decisions based on authority as opposed to fact, and with the technological orientation of contemporary American society.

This view extends to the social as well as to the physical sciences and is reflected in the associations formed to promote the use of social scientific research, associations such as the American Social Science Association founded in the 1860s (Chavkin, 1986), the Social Science Research Council founded in the 1920s, and the National Bureau of Economic Research founded in the 1930s (Goldsmith, 1983), to name a few early examples.

After World War II, there was an explosion of research and development activities in the United States (Rich & Goldsmith, 1982). The expansion of efforts to use research to meet pressing economic and social ends has been so rapid, that since the 1960s we have begun to view science policy and social policy as inextricably linked, with social policy amenable to the laws of scientific inquiry and with science policy having a direct and fundamental influence on the quality of our lives (Rich & Oh, 1994, p. 70). Not surprisingly, knowledge utilization is a major concern of federal agencies concerned with mental health-related issues (cf. Fuller, Caldwell, & Allen, 1992; National Institute of Mental Health, 1992).

Psychology, one of the most recently developed social sciences, has long been dedicated to the use of empirically derived knowledge to improve mental health

KATHLEEN WELLS • Mandel School of Applied Social Sciences, Case Western Reserve University, Cleveland, Ohio 44106-7123.

practice (Meehl, 1997) and mental health policy (cf. Melton, 1987). Indeed, the primary training model for clinical psychology—the scientist-practitioner model—was developed to facilitate such use (Raimy, 1950). And over the past 30 years, empirically derived knowledge has been used to improve some psychotherapeutic practices for children (Kazdin, 1988) and adults (Roth & Fonagy, 1996) and to shape mental health policies (National Institute of Mental Health, 1992).

However, the use of research to improve mental health practice and policy also has been problematic. There are three difficulties. First, some empirically based knowledge that is needed is not yet developed. For example, child mental health policy, as expressed by the Children's and Communities' Mental Health Systems Improvement Act (Public Law No. 102-321), requires states to plan a system of care, involving all the service systems in which children are identified or served; to develop mechanisms through which services to such children are coordinated; and to provide a range of services that are individualized, family-focused, and culturally competent (Rog, 1995). The legislation provides funding for community-based services such as day treatment, outpatient treatment, therapeutic foster or group care, respite care, and emergency mental health services. Yet, we lack evidence for the effectiveness of these services (Bickman, 1996).

Second, some empirically derived knowledge that is known is unused. Practicing clinicians (Cohen, 1985; Morrow-Bradley, Elliott, & Phillips, 1986) and policy makers (Melton, 1987) rarely use research.* For example, despite considerable evidence that multisystemic therapy using a family preservation model of service delivery is more effective than usual juvenile justice system services for adolescents with serious antisocial problems, such approaches have not been implemented widely (Melton, 1997).

The third difficulty is that some empirically derived knowledge is used inappropriately. For example, changes in the health care delivery system and the pressures to identify cost-effective mental health services with which these changes are associated have led to use of research to identify forms of therapy that payers of services will reimburse (Roth & Fonagy, 1996). Yet, translating psychotherapy research into guidelines for community practice is a complex task. Premature use of research findings may dampen clinical creativity, penalize under-researched therapies, promote a foreclosure on the empirical examination of some questions (Roth & Fonagy, 1996), or promote poor practice.

PURPOSES OF CHAPTER

Within the field of pediatric and clinical child psychology, then, the task is not merely to enhance the use of research but rather to promote its appropriate use. This chapter is directed toward that end. Its specific purposes are to examine the

*Moreover, forces within psychology such as the development of the professional school movement (Stricker & Cummings, 1992), which resulted in the development of the PsyD degree, and the American Psychological Society, which formed to emphasize the science over the professional practice of psychology, are widening the gap between research and practice (Stricker, 1997). The more that scientists and practitioners receive different training, "belong to different organizations, attend different meetings, and read different journals, the more separable the paradigms [of science and of practice] will seem, and the more incommensurable they will become" (Stricker, 1997, p. 444).

obstacles to the appropriate use of research in mental health practice and policy, to consider some of the larger issues pertaining to the role of information in enhancing mental health policy and practice, and to suggest some strategies for increasing appropriate use of research. It focuses on practices and policies pertaining to children but not exclusively so. The chapter is based on an examination of the literature on research utilization that spans several social scientific disciplines and professions and on personal experience in applied child mental health research.

DEFINITIONS

Before examining the obstacles to appropriate use of research, it is necessary to consider definitions of research, of users of research, and of use.

Definition of Research

Research is a complex activity involving a set of interrelated tasks: conceptualization, measurement, analysis, and interpretation. Its purposes can include description, hypothesis generation, or hypothesis testing. What is central to each one though is the use of systematic methods to develop knowledge claims based on evidence.

Definition of Users of Research

Potential users of research (Eisele & Gamm, 1981) include public and private agency executives and members of their boards, legislators, members of the judiciary, direct service providers and managers, practitioners who work with individual patients, patients and members of their families, as well as researchers. Although this chapter focuses on two—public policy makers and practitioners—it carries implications for users of research of all types.

Definition of Use

What is used from a single investigation or from a program of research can vary widely (Weiss, 1978; Weiss & Bucuvalas, 1980). Three uses of research are discussed commonly: instrumental, conceptual, and symbolic. Instrumental use is defined as the provision of empirical evidence to select a solution among competing alternatives such as the use of psychotherapy outcome research to determine whether individual psychodynamic or family therapy should be used to treat children with conduct disorder. Conceptual use is the provision of concepts that aid in the conceptualization of a problem such as the use of the primary prevention framework to reconceptualize youth violence as a public health rather than as an individual problem.* Symbolic use is the use of information to legitimate predetermined

*Empirical investigations or research utilization, particularly in public policy making, show the most frequent use of research is conceptual (Weiss, 1978). There is some ambiguity as to this point, however, as most investigations fail to distinguish use of information from its provision, dissemination, and effects (Rich & Oh, 1994).

positions (Beyer & Trice, 1982) such as the use of data on the restricted employment opportunities of teenage mothers by family planning organizations.

OBSTACLES TO RESEARCH UTILIZATION

Irrespective of the definition employed, we lack a theory of research utilization. Moreover, the empirical literature on this topic is at a very preliminary stage of development (Rich & Oh, 1994).^{*} Rather than theoretical or empirical work, the literature is dominated instead by detailed discussions of the obstacles to use of research by academicians who have attempted to influence the research utilization process. Such obstacles include those relating to research, researchers, users of research, and the contexts in which practice or policy is made.

Research

Roth and Fonagy's (1996) analysis of the difficulties applying psychotherapy outcome research to clinical practice, on which the following section depends, illustrates the problems of translating research into mental health practice.

Research Designs

Roth and Fonagy (1996) note that the primary designs used in psychotherapy outcome research—single case studies and randomized controlled trials—limit the extent to which research can be used. Single case studies, which can be implemented in the context of practice and can demonstrate a clinical technique, can neither show its effectiveness nor provide findings that can be generalized across patients and therapists. On the other hand, randomized controlled trials, which can show the efficacy of a standardized technique for a narrowly defined sample, do not provide findings that can be generalized to nonstandardized therapies or to heterogeneous groups of patients. This last observation is an important point. Patients in the community, even those in the same diagnostic category as those examined in a particular investigation, vary widely with respect to factors such as developmental stage, presence of other disorders, social support, and

*Research utilization also has been discussed as part of hypothetical models of knowledge diffusion. Each assumes the existence of a well-defined innovation and as a result has somewhat limited applicability to the discussion here. It is useful to note, though, that Kirk (1979) conceptualizes these models as falling into one of three types: the research development model, the social interactional model, and the problem-solver model. The research development model is a systematic process for "bridging the gap between research and application" of a human service innovation (Rothman, 1978, p. 84). In this model, dissemination of research occurs after a lengthy research and development process in which basic research precedes applied research and applied research precedes development and dissemination of prototypes. The social interactional model is based on the presumption that interpersonal relationships and the social networks of users of research in particular affect the diffusion of an existing human service innovation. The problem-solver model is a user-oriented model of dissemination and utilization. This last model assumes that users of research are at the center of the research and development process. It is compatible with participatory action models of research (Allen et al., 1994; Chesler, 1991) that emphasize the role of service recipients in the design, conduct, and dissemination of a study and with models to aid practitioners in the use research relevant to their practice (Mullen, 1978).

readiness to change, any one of which may exert a powerful influence on the success of treatment.

Concepts Examined

The concepts examined in psychotherapy outcome research also limit its use. The outcomes studied, though linked logically and closely to the interventions provided, do not always reflect the broad changes in functioning desired by patients and therapists alike. Moreover, the timing of the collection of information on patient functioning may be dictated by factors other than a detailed knowledge of the natural history of the disorder under study. Without knowledge of how long an untreated disorder lasts, the way in which the symptoms with which it is associated change over time, and the timing of spontaneous remission and relapses, it is difficult to select meaningful points of time at which to assess the efficacy of a clinical technique.

Data Analytic Approaches

The data analytic approaches employed in psychotherapy outcome research further compromise the utility of such work for practitioners. The data analytic approaches typically employed, which emphasize the identification of statistically significant differences between experimental and control groups, fail to speak to the clinical significance of findings for the range of patients studied, let alone for those seen in community practice.

Character of Findings

Moreover, findings from psychotherapy outcome research can be inconclusive. Psychotherapy outcome research for child treatments is particularly difficult to interpret and use because children, their parents, and their teachers disagree as to children's symptoms and their severity (Kolko & Kazdin, 1993). In a broad way, the problems posed by the effort to use knowledge gleaned from research to improve clinical practice is the problem of trying to achieve internal and external validity in the same investigation (Roth & Fonagy, 1996).

Dissemination of Findings

In addition, problems in the presentation and dissemination of research have been noted. For example, descriptions of empirical investigations may be difficult to follow and dissemination of findings may be lacking or untimely (Eisele & Gamm, 1981).

Interestingly, academicians express the same problem, although stated in somewhat differing forms, in their analysis of the obstacles to use of research to inform public policy. For example, it has been noted that questions examined in research are stated either too specifically or too generally to be useful to policy makers (Corwin & Louis, 1982); that adequate samples are not employed, especially for investigations of social problems involving units other than individuals (Merton, 1949); that studies depend on variables that cannot be manipulated so that their

relevance to action is limited (Beyer & Trice, 1982); that findings are stated in terms of probabilities (Rosen, 1983) or in atheoretical terms (Weiss, 1978); and as a result are difficult to translate into action or to generalize from one situation to another. Therefore, it is not surprising that the few empirical investigations of research utilization we have show that use increases when research is perceived to be of high quality and to carry specific implications for action (cf. Weiss & Bucuvalas, 1980).

Researchers and Users of Research

The roles of research scientist and of public policy maker or practitioner each pose obstacles to the translation of research into practice and policy. For example, researchers are responsible for dispassionately examining questions derived at least in some instances from existing gaps in knowledge, evaluating the limitations as well as the strengths of their investigations, identifying the implications of their work for future research, and publishing their work in scholarly journals for critical evaluation by the scientific community.

As a result, researchers may lack understanding of the complexities of mental health practice or of public policy making; they may resist speculating as to significance of their investigations for concrete situations in the future; they may be reluctant to disseminate the results of their work beyond publication in scholarly journals (Millman, Samet, Shaw, & Braden, 1990); and, in a broad way, they may be uncomfortable combining the role of scientist with the role of practitioner or policy maker, as the requirements for objectivity and careful qualification in the former appear to contradict the requirements for certainty and commitment in the latter.

Users of research such as mental health practitioners and policy makers, on the other hand, are responsible for implementing and defending actions taken in a particular case. As such, they must attend to the evolving features of an individual case, may rely on the opinion of individuals with experience with such cases, and may have insufficient time to peruse the empirically based literature for consensually validated knowledge. Moreover, they may lack the research skills to evaluate critically the empirically based knowledge they obtain, the motivation to change their behavior when such knowledge contraindicates their foundational beliefs and practices (cf. Fuller et al., 1992; Meehl, 1997), or the freedom, due to organizational and other constraints, to change their behavior (Melton, 1997).

For example, Melton (1997) notes an instance of juvenile justice system administrators failing to adopt an innovation they knew to be superior to standard practice because their agencies lacked the capacities to implement new services and were rewarded financially for the provision of existing ones.

Lest the differences between the roles of researcher and user of research be cast too starkly: Researchers are not immune to uncritical thinking (Gross, Levitt, & Lewis, 1996). The National Institute of Mental Health's (1992) report on mental health service system research calls attention to the "deadly effect" ideology can and has had on conceptual thinking of researchers and administrators alike.

Organizational Contexts

The differences in roles of researchers and users of research are exaggerated by differences in the cultures and reward systems of the environments in which they

work. Weiss' (1978) evaluation of the culture of public policy making at the federal level illustrates this point:

Policy makers for their part are interested not only in the application of research evidence to public decisions but also in representing interests and values, reconciling differences, and reaching compromises that maintain the stability of the system. Theirs is political rationality rather than scientific rationality. (p. 61)

ROLE OF RESEARCH TO ENHANCE POLICY AND PRACTICE

Given the complex array of obstacles to using research to enhance mental health practice and policy, it is not surprising that we have lost confidence about "the association of science with unalloyed progress, which many believed was true at one time, ... [or that we have seen an] ominous decline in support of young investigators" (Schneider, 1996, p. 717).

This loss of confidence raises significant questions regarding knowledge development and use in society. These questions include, for example: What should be studied? What role should research play in the development and promulgation of practice and policy? What level of certainty is required for action? Formulation of sensible strategies to enhance appropriate use of research for mental health policy and practice requires a consideration of such questions.

Limitations on Utility of Research

The extent to which social scientific research can contribute to social problem solving may be more limited than is supposed generally. Following Lindblom and Cohen's (1979) propositions, it is clear that empirically derived knowledge is only one source of information for social problem solving and that values, ordinary knowledge, and social action are other important sources of information in this regard.

Indeed, the potential of social scientific research is limited by the intricacy and fluidity of social life. Such complexity makes it difficult to produce generalizable knowledge (Harris, 1986) and, more to the point, to know in advance which problems are context-bound and which are not. Indeed, questions continue to be raised as to whether psychology can ever be an accumulative science (Koch, 1981; Toulmin, 1972, cited in Schneider, 1996).

Irrespective of one's views on this issue, Lindblom and Cohen's (1979) discussion underscores the need for modesty regarding the potential of research to contribute to mental health policy or practice; for consideration of the nature of that influence compared to that provided by values, ordinary knowledge, and social action; and for delineation of the constraints on external validity of any investigation.

Nature of the Contribution To Be Made

The failure of practitioners and policy makers to use research also has been linked to the extreme specialization that has occurred within the social sciences [see Bevan's (1991) discussion of this for psychology]. Such specialization has worked to denigrate the discussion of general ideas and to disconnect considera-

tion of the ways in which specific areas of study are connected to broad cultural problems. This is not a new condition, however. Sixty years ago, Lynd (1939) made a similar observation.

Therefore, one approach to enhancing the use of research in mental health practice and policy is to recast mental health research so that it is linked explicitly to broad cultural problems and to involve work of investigators from all relevant disciplines. As Bevan (1991) has observed: "Specialization requires integration; the two are not mutually exclusive ... [we must] transcend disciplinary boundaries altogether" (p. 481).

This perspective could work to minimize much of the discussion as to whether a particular investigation is "relevant" or not. Experience shows that broad cultural problems such as aggression, alienation, or impoverishment are relatively enduring. Careful selection of aspects of those that could be examined productively within an empirical framework and a long-term commitment on the part of investigators and funders of research to their study would go a long way toward the development of usable knowledge.

STRATEGIES FOR ENHANCING APPROPRIATE USE OF RESEARCH

Specific strategies to enhance appropriate use of research flow from a consideration of the obstacles to such use and of the potential for empirically derived knowledge to enhance mental health practice and policy. Strategies pertaining to doctoral-level training, research development and implementation, and research presentation and dissemination are noted below.

Doctoral Training

By and large, doctoral-level training programs in psychology, as well as programs in other disciplines and human service professions, socialize students into an approach to social science that is ahistorical (focused on prevailing assumptions rather than on their evolution over time) and technocratic (focused on techniques of measurement and analysis rather than on ideas and their relationships). Moreover, training in rather narrowly defined substantive areas is expected to stand for a broad orientation to mental health problems of the culture; training in research design and statistics is expected to stand for a broad orientation to philosophy of science. Serious attention to the approaches taken by a range of social scientific disciplines to a problem under study is lacking. This socialization limits students' abilities to conduct research that can be used because it is too narrow.

As a result, I recommend we alter doctoral-level training programs so that they are both more interdisciplinary and problem focused than they are at the present time. This might involve changing the curriculum to emphasize:

1. The study of the history of science, with particular attention to current controversies regarding the intellectual structure of a discipline and its rhetoric of inquiry.
2. The evolution of mental health practices and policies over at least the past 100 years, with particular attention to the ways in which concepts have developed, been abandoned, and or revised over time.

3. The examination of the ways in which a particular substantive area is defined and approached by a wide range of disciplines, when specialization does occur.

For example, one component of a program might focus on neglect of children. Within this area, students would examine approaches to the definition and treatment of neglect and how they have evolved over time; ways in which neglect is defined and assessed currently within the context of law, medicine, and psychological practice; prevailing macro- (such as economic) and micro-level (such as psychological) analyses of the causes and solutions to the problem; and current controversies and gaps in knowledge.

In this way, students could be helped to identify and to grapple not only with an enduring cultural problem but also to bridge the gap between research and practice by emphasizing the critical elements common to both activities: thinking critically and theoretically about complex problems.

Within this broad framework, the scientist-practitioner model of graduate training in clinical psychology would fit naturally. It needs to be implemented, however, more fully than it has in the past, with due attention to both practice and to science as well as to their integration (Rodnick, 1966). As a result, I also recommend we alter doctoral-level training programs so that they focus directly on the links among theory, practice, research, and research utilization. This might involve changing the curriculum to emphasize:

1. The examination of ways in which paradigm shifts, emerging concepts, and empirical findings have affected specific practices.
2. The identification of practices that have and have not been validated empirically, emphasizing that “reduction in uncertainty gained through research should not be compared with an ideal but unlikely state of absolute knowledge, but with an absolute state of ignorance regarding the phenomena” (Rosen, 1983, p. 11).

It might also involve:

1. The development of practicums and opportunities for consultation that link practice and research.
2. The development of courses that examine closely the problems and possibilities of research utilization.

With respect to the last two suggestions, programs might adopt the case method, a pedagogical approach used widely in schools of business, to study examples of successful and unsuccessful research utilization and to provide a forum in which students and faculty providing research consultation could seek advice. In this way, students could be helped to understand the complexities of research utilization as well as the uses and misuses to which research can be put in policy and practice.

Research Design and Implementation

Efforts to conduct investigations that have high internal validity often work to produce investigations of limited external validity. Limitations on external validity

oftentimes curtail the utility of research for mental health practice and policy making.

As a result, I recommend we promote development of programs of research rather than design of single investigations. This would allow the complexity of phenomena to be captured to a greater degree than it is at the present time. Research program development might include several design and implementation strategies to enhance utilization of research such as the following:

1. Identify questions to study, based on an assessment of the degree to which the program of research will be able to contribute to knowledge of an enduring cultural problem.
2. Cast questions based on detailed knowledge of the mental health practices or policies to which such questions are linked logically.
3. Employ, where possible, an interdisciplinary approach to study of the problems under investigation.
4. Examine the full range of phenomena under study to promote explanations of both typical and of exceptional cases.
5. Employ diverse approaches to study of phenomena, including the use of qualitative methods, so that the complexity of phenomena can be captured.
6. Have senior-level investigators immerse themselves in the data collection phase of projects, so that the most experienced individuals are exposed to the reality of the problems under study.
7. Qualify rigorously study findings but also speculate freely as to their relationship to general ideas.

These strategies hold the potential to yield empirically derived knowledge that is both well-developed and related to policy or practice.

Research Dissemination

By and large, users of research, other than researchers, are uncertain as to the standards by which to judge the quality of research on which its significance for practice or policy ultimately depends. As a result, I recommend we develop mechanisms other than publication in scholarly journals for the dissemination of research to the practice and policy communities. Strategies such as the following ones might include:

1. Promotion of efforts to integrate bodies of knowledge pertinent to specific problems in mental health practice or policy.
2. Review of these integrations by scholars in order to develop a consensus regarding knowledge in a particular area.
3. Translation of such consensually validated knowledge into specific implications for policy and practice.
4. Review of implications by a wide range of users of research for their significance for recommendations for improving policy and practice.

While strategies such as these have been implemented by professional associations such as the American Psychological Association and federal agencies, their replication by teams of investigators in the context of their communities would promote appropriate use of research at the local level and build a framework in

which academicians, practitioners, and policy makers could work together, over the long haul, on difficult problems.

INSTITUTIONAL DEVELOPMENT

Such strategies require time to implement as well as, one might add, cooperation and restraint. As such, they are at variance with the structure of most academic departments and schools, which emphasize empirical work over knowledge integration and which reward individual effort over interdisciplinary collaboration. They also are at variance with the structure of most practice and policy-making groups, which emphasize promotion of solutions rather than the analysis of problems. Moreover, agencies typically lack the capacity to implement reform.

Thus, serious consideration of the strategies delineated above would require changes in institutions to provide the structure and support for the production of usable knowledge.

Within academic departments, such change might take the following forms:

1. Hire and promote faculty who are experienced clinicians as well as those who are experienced researchers to advance the integration of research and practice.
2. Develop incentives for faculty to contribute to the dissemination of research and knowledge, broadly defined, particularly within their local communities.
3. Develop programs whose purpose is to forge links between faculty and practitioners and policy makers.

For example, centers could be developed, such as the Family Services Research Center at the Medical University of South Carolina (Henggeler, Melton, Brondino, Scherer, & Hanley, 1997) dedicated to the development, test, specification (through development of treatment manuals), and dissemination of specific psychotherapeutic interventions for children and their families. New institutional arrangements between academic departments or schools and other agencies in the community, including state departments of mental health, local mental health boards, public or private agencies, or individual programs could be forged. One such example is the Cuyahoga County Community Mental Health Research Institute, a collaborative project of the Mandel School of Applied Social Sciences and the Cuyahoga County Community Mental Health Board, which has developed a research program central to the board's mission and developed policies and procedures to ensure use of research produced (Biegel, Wells, Johnsen, & Dowhower, 1996). Within public agencies, such change might involve:

1. Promotion of leaders willing to evaluate dispassionately their own practices.
2. Development of organizational units with responsibility for research utilization.

A review of the obstacles to appropriate use of research in mental health practice and policy making and the development of strategies to overcome them can be encapsulated, by way of conclusion, in following broad directives. First, we need to emphasize development of well-conceptualized programs of research

because they hold more potential to inform the public debate than do single investigations. Second, we need to link these programs to ongoing cultural problems because such problems require sustained study and demand attention from the scholarly community. Third, we need to endorse the scientist-practitioner model as the primary training model in clinical psychology because it is the foundation on which to build opportunities for both students and faculty to promote appropriate use of research. Fourth, we need to develop institutional structures to support research utilization, both within and outside academia, because it is a time-consuming, costly, and complex activity that requires the talents of researchers, practitioners, and policy makers as well as consumers and payers of services, to name a few.

ACKNOWLEDGMENTS. The author thanks Drs. Dennis Drotar and Sandra Russ for their helpful comments on a prior draft of this chapter.

REFERENCES

- Allen, D., Gilchrist, L., Brown, L., Cox, G., Semke, J., Thomas, M., & Perry, R. (1994). One system, many perspectives: Stakeholders and mental health system evaluation. *Evaluation and Program Planning*, 17(1), 47-51.
- Bevan, W. (1991). Contemporary psychology: A tour inside the onion. *American Psychologist*, 46(5), 475-483.
- Beyer, J., & Trice, H. (1982). The utilization process: A conceptual framework and thesis of empirical findings. *Administrative Science Quarterly*, 27(4), 591-622.
- Bickman, L. (1996). A continuum of care: More is not always better. *American Psychologist*, 51(7), 689-701.
- Biegel, D., Wells, K., Johnsen, J., & Dowhower, D. (1996). *Facilitators of and barriers to research utilization: A case study*. Final report to the Cuyahoga County Community Mental Health Board, Cleveland, OH.
- Chavkin, N. (1986). The practice-research relationship: An organizational link. *Social Science Review*, 60(2), 241-250.
- Chesler, M. (1991). Participatory action research with self-help groups: An alternative paradigm for inquiry and action. *American Journal of Community Psychology*, 19(5), 757-768.
- Cohen, L. (1985). Research utilization in community mental health. Some comments on O'Donohue, Curtis, and Fischer. *Professional Psychology, Research, and Practice*, 16(5), 719-722.
- Corwin, R., & Louis, K. (1982). Organizational barriers to the utilization of research. *Administrative Science Quarterly*, 27(4), 623-640.
- Ersele, F., & Gamm, L. (1981). Research utilization: Reaching decision makers. *Children and Youth Services Review*, 3(4), 291-303.
- Fuller, R., Caldwell, R., & Allen, J. (1992). Utilizing alcoholism treatment research findings in clinical practice. In R. Watson (Ed.), *Drug and alcohol abuse reviews*, Vol. 3: *Alcohol abuse treatment* (pp. 187-202). Totowa, NJ: Humana Press.
- Goldsmith, N. (1983). The study of research utilization in mental health policy. *Dissertation International*, 44(5-B) 1637.
- Gross, P., Levitt, N., & Lewis, M. (Eds.) (1996). *The flight from science and reason*. New York: The New York Academy of Sciences.
- Harris, J. (Ed.) (1986) *Child psychology in action*. Cambridge, MA: Brookline Books.
- Henggeler, S., Melton, G., Brondino, J., Scherer, D., & Hanley, J. (1997). Multi systemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology*, 65, 821-833.
- Kazdin, A. (1988). Child psychotherapy: Developing and identifying effective treatments. Oxford, England: Pergamon Press.
- Kirk, S. (1979). Understanding research utilization in social work. In A. Rubin & A. Rosenblatt (Eds.), *Sourcebook on research utilization: A compilation of papers presented for discussion at the*

- Conference on Research Utilization in Social Work Education* (pp. 3–15). New York: Council on Social Work Education.
- Koch, S. (1981). The nature and limits of psychological knowledge: Lessons of a century qua “science.” *American Psychologist*, 36, 257–259.
- Kolko, D., & Kazdin, A. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent, and teacher reports. *Journal of Child Psychology and Psychiatry*, 34, 991–1006.
- Lindblom, C., & Cohen, D. (1979). Usable knowledge: Social science and social problem solving. New Haven, CT: Yale University Press.
- Lynd, R. (1939). *Knowledge for what? The place of social science in American culture*. Princeton, NJ: Princeton University Press.
- Lynn, L. (1978). *Knowledge and policy: The uncertain connection*. Washington, DC: National Academy of Sciences.
- Meehl, P. (1997). Credentialized persons, credentialized knowledge. *Clinical Psychology: Science and Practice*, 4(2), 91–98.
- Melton, G. (Ed.). (1987). *Reforming the law: Impact of child development research*. New York: Guilford Press.
- Melton, G. (1997). Why don’t the knuckleheads use common sense? In S. Henggeler & S. Santos (Eds.), *Innovative approaches for difficult-to-treat populations* (pp. 351–370). Washington, DC: American Psychiatric Press.
- Merton, R. (1949). The role of applied social science research in the formation of policy: A research memorandum. *Philosophy of Science*, 16(3), 161–181.
- Millman, J., Samet, S., Shaw, J., & Braden, M. (1990). The dissemination of psychological research. *American Psychologist*, 45(5), 668–669.
- Morrow-Bradley, C., & Elliott, R. (1986). Utilization of psychotherapy research by practicing psychotherapist. Special issue: Psychotherapy research. *American Psychologist*, 41(2), 188–197.
- Mullen, E. (1978). The construction of personal models for effective practice: A method for utilizing research findings to guide social interventions. *Journal of Social Service Research*, 2(1), 45–63.
- National Institutes of Health and Mental Health. (1999). *Bridging science and service: A report by the National Advisory Mental Health Council’s clinical treatment and services research workgroup* (NIH: Publication No. 99–4353). Rockville, MD.: Author.
- National Institute of Mental Health. (1992). Caring for people with severe mental disorders: A national plan to improve services: Research resources. *Schizophrenia Bulletin*, 18(4), 669–696.
- Public Law No. 102-321. (1992). ADAMHA Reorganization Act of 1992, 201(2), 106 Stat. 378.
- Raimy, V. (1950). *Training in clinical psychology*. New York: Prentice-Hall.
- Rich, R., & Goldsmith, N. (1982). The management and utilization of research and development. *Knowledge: Creation, Diffusion, and Utilization*, 3, 415–436.
- Rich, R., & Oh, C. (1994). The utilization of policy research. In S. Nagel (Ed.), *Encyclopedia of policy studies* (2nd ed, pp. 69–92). New York: Dekker.
- Rodnick, E. H. (1966). Comments on the “Boulder” model. In E. L. Hoch, A. Ross, & C. Winder (Eds.), *Professional preparation of clinical psychologists* (pp. 21–23). Washington, DC: American Psychological Association.
- Rog, D. (1995). The status of children’s mental health services: An overview. In L. Bickman & D. Rog (Eds.), *Children’s mental health services* (pp. 1–18). Thousand Oaks, CA: Sage.
- Rosen, A. (1983). Barriers to utilization of research by social work practitioners. *Journal of Social Service Research*, 6(3–4), 11–15.
- Roth, A., & Fonagy, P. (Eds.). (1996). *What works for whom? A critical review of psychotherapy research*. New York: Guilford Press.
- Rothman, J. (1978). Conversion and design in the research utilization process. *Journal of Social Service Research*, 2(1), 117–131.
- Schneider, S. (1996). Random thoughts on leaving the fray. *American Psychologist*, 51(7), 715–721.
- Stricker, G. (1997). Are science and practice commensurable? *American Psychologist*, 52(4), 442–448.
- Stricker, G., & Cummings, N. (1992). The professional school movement. In D. Freedheim (Ed.), *History of psychotherapy: A century of change* (pp. 801–828). Washington, DC: American Psychological Association.
- Weiss, C. (1978). Broadening the concept of research utilization. *Sociological Symposium*, 21, 20–33.
- Weiss, C., & Bucuvalas, M. (1980, April). Truth tests and utility tests: Decision makers’ frames of reference for social science research. *American Sociological Review*, 45, 302–313.

Index

- Abstract, in grant application, 228
- Acquired Immune Deficiency Syndrome (AIDS), children's intuitive theories of, 157, 152; *see also* Human Immunodeficiency Virus
- Adolescent risk behavior, 37, 41
as example of evaluating structural model, 37
test of mediation model for, 441
- Advisory council, in National Institutes of Health (NIH) research review, 223–224
- African-American population, anxiety disorders in, 173
- African Self-Consciousness Scale (ASCS), 169
- American Association of Behavioral Therapy, 332
- American Indian children, research with, 175
- American Journal of Orthopsychiatry, 351
- American Psychiatric Association, research funding and resource manual of, 215
- American Psychological Association, 412, 540
composition of ethnic minority professionals in, 182
- American Psychologist, 349
- American Social Science Association, 531
- AMOS, use in evaluating structural models, 37
- Analysis of change, 53–54
advantages of multi-level approach in, 69–71
incremental view of, 53
process view of, 54
use of difference scores in, 53–54
- Annual Register of Grant Support, 213
- Armington Foundation, 208
- Assessing Coordinated Care Study (ACCS), 476–477
- Assessment
behavioral analytic model in, 129–130
conceptual limitations of, 120–123
contextual approach to, 127–133
diary methods in, 133–136
family level methods of, 126
functional outcomes in, 120–124
global versus contextual methods of, 120–122
methodological limitations of, 124–127
- Association for Advancement of Behavior Therapy (AABT), 430
- Attention Deficit Hyperactivity Disorder (ADHD), impact of severity on sampling, 83
- Attrition bias, 80–82
in sample selection and recruitment, 80–83
strategies to reduce, 89
statistical methods, 491–492
- Attrition, of investigators in multi-institutional research, 295
- Authorship, *see also* Editing manuscripts; Reviewing manuscripts; Writing research articles
strategies of in multi-institutional research, 299–300
- Background and significance, in research grant proposal, 231–232
- Baltimore City Public School System, 10
- Bayley Scales of Development, 179–181
- Behavior analytic model, 129–133, 157
in assessing stress and coping, 129–130
instrument development in, 131
item development in, 130
response enumeration in, 130–131
response evaluation in, 131
- Role Play Inventory of Situations and Coping Strategies (RISCS)
as example of, 131–133
situational analysis in, 130
- Behavior Therapy, 361, 380
- Bentley's Comparative Fit Index, 38; *see also* Fit index
- Between group studies, of ethnic minorities, , 170–172
- Biographical sketch, in research grant proposal, 229–230
- Bone marrow aspiration, 498
- Budgets
managing in research, 256–257
in applications to foundations, 205
in applications to government agencies, 230

- Buffering hypothesis, 27
 Buffering models, 24–25, 27
 Bureau of Health Care Delivery and Assistance, 216
 Maternal and Child Health Branch of, 216, 521
- Case control design, 15–16
 Case reports, *see also* Case studies and series
 assessment occasions in, 434–436
 evaluation of, 433–446
 example of, 433–434
 number of subjects in evaluating, 440
 heterogeneity of subjects in evaluating, 440
 role of past course in evaluating, 437–438
 stability of therapeutic effects in evaluating, 439–440
 types of data in, 434–436
 types of effect in, 438–39
 Case studies and series, 429–447; *see also* Case reports
 clinical effects in, 442–443
 evaluation of, 433–442
 example of case report in, 433–436
 group and single-subject designs in, 4, 430–432
 guidelines for publishing in *Journal of Pediatric Psychology*, 430
 single subject research in, 431
 stability of therapeutic effect in, 430–440
 training in methods of, 443–445
 types of, 432–433
 types of effects in, 438–439
- Case Western Reserve University, research training in, 208–209
 Centers for Disease Control, 216
 Child Adult Medical Procedure Scale (CAMPIS), 499, 504, 508
 Child and Adolescent Service System Program (CASSP), 469
 Child Behavior Check List (CBCL), 15, 80, 88, 100, 124
 Child Behavioral Therapy, 351
 Child development, 351
 studies employing race comparative designs in, 171
 Child Health Questionnaire, 201
 Childhood chronic illness
 measurement of stress and coping in, 121–122
 relationship to maternal mental health, 41–42
 risk for emotional disorder in, 123
 test of mediation model in, 41–42
 Child mental health system, conducting research in, 274–277
 Child welfare settings, conducting research in, 272–273
 Children's Cancer Group (CCG), 287, 521
 in multi-institutional collaborative research, 287
- Children's and Communities' Mental Health Systems Improvement Act (PL102-321), 532
 Children's Depression Inventory (CDI), 120, 122, 124, 362
 Children's Depression Rating Scale (CDRS), 102
 as example of short form of existing measure, 103
- Children's Health Care, 351
 Children's Hospital of Pennsylvania, 511–525
 Psychosocial Services Program in Division of Oncology in
- Children's Services, Social Policy, Research and Practice, 351
 Classroom social ecology, 6–8
 Cleveland Foundation, 196
 Clinical Child and Family Psychology Review, 380
 Clinical Psychology Review, 380
 Clinical child psychology
 advocacy for research funding in, 240–241
 application of qualitative methods in, 155–161
 journal options in, 350
 Clinical practice, as model for clinical research, 495–498
 Christchurch Health and Development Study, 42
 Chronic illness: *see* Childhood chronic illness
 Chronicity
 role of sampling variation in, 83
 of stressors, 127–128
 Clinical relevance, in mental health services research, 481–482
 Clinical significance, 354
 Cluster sampling, 79
 Cofactors, in sampling bias, 91
 Cognitive Development, manuscript in review, 419
 Cognitive Rating Scale, 101–113
 example of modifying existing scale, 107–113
 example with suicide attempts, 101–113
 Cohen's kappa, 100
 Collaboration, in randomized studies, 281–305
 Collaborative research models, 281–305; *see also* Multi-institutional research collaboration
 Community epidemiology, 6–10
 defining ecological context in, 6–7
 effects of classroom ecology in, 7–8
 effects of neighborhood social ecology in, 8–9
 sampling in, 9–10
 use of weighted samples in, 8–9
 Comparative fit index, 38; *see also* Fit index
 Comparisons with test norms, 38
 Competing continuation application, 225
 Comprehensive Mental Health Services Program for Children, 474
 Conditional agreement, definition of, 115

- Conditional model, in individual growth curve analysis, 66–69
- Conditional probability approach, 115
- Conduct disorder, neuropsychological antecedents of, 31
- Confidentiality
in research, 307–309
in review process, 419
- Confirmatory analysis, 33–34
- Confounders, 27; *see also* Confounding variables
- Confounding variables, 33–34, 45; *see also* Confounders
- Connecticut Longitudinal Study (CLS), 51, 57–67
application of individual growth curve analysis in, 57–66
- Concept mapping, 475
- Confirmability, as criteria of qualitative method, 183
- Consent rate, 81
impact on sampling in published studies, 81
- Consultation, on research grants, 226
- Constructs, 32–35
endogenous, 32–33
exogenous, 32–33
multiple criterion, 35
single criterion, 35
- Context specific measures, in chronic illness, 122
- Continuing education opportunities
in applied research methods, 279
in managing research, 260
- Contextual approach: *see* Assessment
- Continuous Performance Test, 105
- Control group, selection of in sampling, 88
- Convenience sampling, in ethnically diverse samples, 173–174
- Cooperative Study of Sickle Cell Disease (CSSCD), 282, 287–289, 293
- Coping: *see Stress and coping*
assessment of, 129–130
assessment of with Kid Cope, 105–107
application of behavior analytic model in, 129–133
behavioral analytic model in assessment of, 129–133
models of, 23
Role Play Inventory of Situations and Coping Strategies (RISCS),
as measure of, 131–133
- Correlated error term, in testing structural models, 39
- Council of Biology Editors Peer Review Retreat Consensus Group, 410
- Covariance analysis, in testing bias for attrition, 91–92
- Covariates, 45
- Credibility, as criteria in evaluating qualitative studies, 151–155
- Cultural influences, in research with pediatric and clinical child psychology populations, 165–184
- Cultural relevance, 179–182
identification of norm groups to enhance, 181
measures of, 181–182
procedures to enhance, 179–181
- Culturally sensitive measures, 181–182
- Culture, as a variable in research, 168
- Cuyahoga County Community Mental Health Research Institute, 541
- Cystic fibrosis, 86, 506
application of Role Play Inventory of Situations, and Coping Strategies (RISCS) to, 131–133
- influence of mortality on sampling in, 86
problem situations in treatment of, 131–133, 157
- Daily diaries, 123–124, 126, 157; *see also* Daily Phone Diary Procedure
- Daily Phone Diary Procedure (DPD), 135–136; *see also* Daily diaries
as example of context specific assessment method, 135–136
- Data, qualitative, 148–149
- Data analysis, in research proposals to government agencies, 235–236
- Data analytic models, 25–33
confirmatory analysis in, 133
direct effects in, 25–26
endogenous constructs in, 32–33
evaluating adequacy of, 37–39
exogenous constructs in, 32–33
fit index in, 37–39
types of
interaction model, 27–28
measurement model, 33
mediation model, 28–30, 33
structural equation models, 32–36
types of constructs in, 32
- Data analytic procedures, *see also* Growth curve models; Hierarchical linear modeling; Individual growth curve analysis; Structural equation modeling
for assessing sampling bias due to selective attrition, 91–92
for clarifying impact of sampling procedures, 91
for adjusting cofactors, 91
- Data collection
in research with ethnic minority samples, 176–178
cultural barriers in, 176
language barriers in, 176
impact of socioeconomic status in, 176

- Data collection procedures, in research grants, 234–235
- Data management, in research, 254–255
- Decision making
- in anticipating sampling problems, 86–87
 - in mixed methods research designs, 155
 - in research on ethnic minorities, 167–170
- Dependability, as criteria for evaluating qualitative studies, 153
- Developmental epidemiologic framework, in research with children, 3–17
- Developmental epidemiology, efficiency of multi-stage sampling in, 13–15
- Development office, role in proposals to foundations, 195–196
- Diabetes, supportive behaviors in care of, 157
- Diabetes Control and Complications Trial, 282, 287, 291, 292
- Diagnostic and Statistical Manual of Mental Disorders (third edition) (DSM-III), definition of impulsivity in, 104
- Diathesis stress models, 25, 27
- Difference score approach, in analyzing change, 53–54
- Direct effects models, 25
- Directory of Research Grants, 213
- Dissemination of information
- strategies of, 277, 278
 - from research findings, 329, 540–541
- Distal factors, 31
- in model of acute procedural pain, 501
- Division of Research Grants (DRG), 220
- Documentation, as criteria for evaluating qualitative studies, 154
- Duration of condition, impact on sample variation, 83
- Ecological psychology**, as qualitative method, 151–152
- Editing manuscripts, 352–364, 409–423; *see also* Editorial board members; Editorial feedback; Editorial responsibilities; Reviewing manuscripts; Writing research articles
- ethical issues in, 418–420
 - confidentiality of review in, 419
 - ensuring quality and accuracy in, 419
 - managing conflicts of interest in, 419
- Editorial biases, 420
- Editorial board members, roles and responsibilities of, 413–415, 421
- Editorial board structure, 410–413
- Editorial feedback, types of, 418
- Editorial responsibilities, 410–422
- Effect size, 396
- Effective reviews, characteristics of, 415–418
- Effectiveness, 465–467
- Endogenous constructs, 32–33
- End point analysis, in testing biases due to attrition, 92
- Epidemiology, *see also* Developmental epidemiology
- training clinical and pediatric psychologists in, 15–17
- EQS, use of in evaluating structural model, 37
- Ethics, in manuscript review, 418–420
- Ethics in research, 305–324, 316–317, 457–460
- confidentiality, 307–309, 457–458
 - consent forms in, 311–312, 323–324
 - emergency procedures and guidelines in, 322–323
 - future directions in, 319–321
 - in follow-up studies, 315–316
 - informed consent, 310–314
 - in randomized treatment studies, 457–460
 - in psychiatric hospitals, 318
 - with vulnerable child and adolescent populations, 309–310
 - depressed research participants, 309–310
 - suicidal participants, 310
- Etiology of condition, role in sampling variation, 83–84
- Ethnic influences in research, with child clinical and pediatric psychology populations, 165–184
- Ethnic identity, measurement of, 177
- Ethnicity, terminology in, 168–169
- Evaluating mental health services, 462–484; *see also* Mental health services research
- Assessing Coordinated Care Study as example of, 476–477
 - Child and Adolescent Service Program in, 469
 - clinical relevance of, 481–483
 - concept mapping in, 479
 - Diagnostic Specific Treatment Improvement Approach in, 481
 - effectiveness in, 465–467
 - efficacy in, 465–467
 - fidelity of treatment in, 473–474
 - Fort Bragg Evaluation Project as model of, 467–484
 - generic services improvement approach in, 481
 - implementing monitoring of, 472–474
 - interpreting null findings in, 480–481
 - perceptions of multiple stake-holders in, 478–479
 - program theory in, 469–470
 - randomized controlled trials in, 466
 - recruitment and eligibility in, 474–475
 - referred populations in, 477–478
 - Stark County Evaluation Program in, 468, 471, 478–479
 - research in, 467–481
 - theories of implementation in, 474
 - theory based measurement in, 470–472

- Evaluation
of case report, 433–442
of meta-analysis, 395–405
of treatment: *see* Randomized treatment studies; Evaluating mental health services
- Exclusionary criteria, making decisions about, 86
- Experimental design and method, in research grants, 233–237
- Exogenous constructs, 32–33, 45
- Explicative research, 492–494
- External validity
in mental health services, 468
in randomized treatment studies, 453–457
- Eysenck Impulsiveness Scale, 105
- Fagan Test of intelligence, 179–180
- Failure to thrive
impact of etiologies on sampling, 83
impact of referral patterns on sampling, 84
- Faking bad, in psychological measures, 98
- Faking good, in psychological measures, 98
- Family demographic characteristics, role in sampling, 85
- Family Environment Scale, 125
- Family level assessment, 126
- Family Life Events Interview, 126
- Family Services Research Center, at University of South Carolina, 541
- Family systems theories, 23–24
- Federal Register, 216
- Fidelity, of treatment implementation, 473–474
- Fit index, 34, 37–39
Bentley's comparative, 38
normed and non-normed, 38
Tucker Lewis Index, 38
chi-square-df ratio, 38
in testing structured models, 34, 37–39
- Fixed effects, in individual growth curve models, 57–66
- Fixed quadratic growth parameter, 62–65
- Florida Conference on Child Health Psychology, 332
- Fort Bragg Evaluation Project, 467–484
- Foundation Directory, 198, 213
- Foundations
available funding from, 196–197
facilitating successful applications to, 198
making contact with, 195–196
preparing budgets for, 204
preparing research proposals for, 191–202
research priorities of, 192–295
review process for, 205–206
feedback in, 205–206, 208
site visit in, 206–207
steps in application to, 194–195
types of, 192
- Funding: *see* Foundations; Research funding; Research grant proposals, to government agencies
- Future directions
in ethical conduct of research, 319–321
in evaluating mental health services, 483–484
in funding research, 240–241
in research on qualitative methods, 159–161
- Gatlinburg Conference on Mental Retardation, 332
- Genentech Foundation for Growth and Development, 198
- Generalizability
of findings in research reviews, 392
of research with children, 87–93
strategies to enhance, 92–93
in qualitative methods, 149
- Goodness of fit, in evaluating growth curve models, 58
- Grants: *see* Foundations; Research funding; Research grant proposals, to government agencies
- Grant Directory, 213
- Grant support, for multi-center research, 300
- Great Lakes Society of Pediatric Psychology, 332
- Grounded theory
application of, 155–156
as qualitative method, 150–151, 155–156
- Growth and Development Study, 92
as example of multi-site study, 92
- Growth curve models 51, 57–62; *see also* Individual growth curves analysis
conditional model, 57
examples of, 51
unconditional model, 57–66
- Harter Self-Perception Scale, 362
- Head Start, 11–12
- Health psychology, 351
- Hierarchical linear model, 56
in individual growth curve analysis, 56
- Home Observation For Measurement of the Environment Scale, 182
application to Ugandan infants, 182
- Human Immunodeficiency Virus (HIV), 172, 513; *see also* Acquired Immune Deficiency Syndrome
in Ugandan infants, 178–182
- Human subjects
in research proposals to National Institute of Health (NIH), 236–237
review boards, 237, 271
- Hypotheses, in research proposals, 234
- Implementation
challenges in research with ethnic minority patients, 173–174
of cross-cultural research, example from studies of Ugandan infants, 178–182
of mental health services research, 472–474

- Implementation (*cont.*)
 of research with pediatric and child clinical populations, 243–266
 of research in schools, 269–271
- Impulsive Behavior Check List**, as example of self-report measure, 104–105
- Impulsive behavior**: *see* Impulsivity
- Impulsivity**
 measure of in suicide attemptors, 104–105
 definition of in Diagnostic and Statistical Manual of Mental Disorders, 104
- Incidence**, as selection criteria for, 86
- Inclusionary criteria**, 85–87
 decisions concerning, 86–87
- Incomplete Sentence Blank**, 102
- Incremental view**, of change, 83–84
- Individual growth curve analysis**, 51–78
- Individual Growth Curves (IGC)**, 51–57
 estimation of, 56
 hierarchical linear models in, 56–61
 mathematical model of, 55–56
 multilevel models in, 56–57
 use of ordinary least squares estimates in, 57
 use of weighted least squares in, 57
- Inclusion of children**, in proposals to NIH, 237
- Inclusion of minority participants**, in proposals to NIH, 236–267
- Inclusion of women**, in proposals to NIH, 236–237
- Infant Health and Development Program**, 85, 92
 as example of multi-site study, 92
- Institutional review boards (IRBs)**, 237, 306–307; *see also* Human subjects, 237, 306–307
- Integrating research and policy**
 in mental health research, 529–574
 definitions of, 533–534
 institutional development in, 541–542
 obstacles to research utilization in, 534–536
 role of research in, 537–538
 role of Children's and Communities' Mental Health Systems Improvement Act in (PL102–321), 532
 strategies to enhance, 536–541
 doctoral training in, 536–537
 research design, 540–541
- Integrating research and practice**
 acute procedural pain in children as a model of, 498–505
 clinical practice as a model of, 495–498
 explicative research in, 492–494
 empirically derived treatment and research in, 502–505
 in pediatric hospital, 511–526
 in pediatric psychology, 491–509
 principles of social ecology in, 514–516
- Psychosocial Services Program**, Division of Oncology, Children's Hospital of Philadelphia as example of, 511, 516–524
- Integrating research and practice (*cont.*)**
 role of pediatrics and psychology in, 511–514
 recommendations to facilitate, 505–509
 treatment research in, 492–494
- Interaction effects**, 28–30
 examples of, 28–30
- Interaction model**, 27–30
- Internal validity**, in conduct of randomized trials, 448–453
- Johns Hopkins University**, 16, 480
 Prevention Intervention Center, 3–6, 10, 14
- Journal of Abnormal Psychology**, 351
- Journal of Abnormal Child Psychology**, 351–352, 412
- Journal of the American Academy of Child and Adolescent Psychiatry**, 351
- Journal of Black Studies**, 167
- Journal of Child Psychology and Psychiatry**, 351
- Journal of Consulting and Clinical Psychology**, 351
 section on recruiting and maintaining minorities in research, 175
 section on case studies, 350
- Journal of Clinical Child Psychology**, 350
 consent rates in published studies in, 81
 consent and assent in published studies in, 321
 editorial board of, 410
 recent articles in, 245
 review articles in, 380
 qualitative methods in, 147
 studies funded by foundations in, 191
 studies of ethnic minority children in, 183
- Journal of Developmental and Behavioral Pediatrics**, 351, 380
 blinding reviews in, 422
- Journal of Pediatric Psychology**, 351
 consent and assent in published studies in, 322
 consent rates in published studies in, 81
 editorial board of, 410
 explicative research in, 492
 qualitative methods in, 147
 recent articles in, 245
 review articles in, 380
 reviewers' recommendation form for, 352
 special issues of, 411
 studies of ethnic minority children in, 116
 timeliness of reviews for, 417
 treatment research in, 492
- Journals in Psychology**, 350
- Juvenile migraine headaches**, 199–201
- Kansas Conference On Clinical Child Psychology**, 332
- Kid Cope**, 105–107
 as measure of coping, 105–107
 development of, 105–107
 factor analysis of, 106–107
 reliability of, 106–107

- Language barriers, to data collection with ethnically diverse samples, 176
- Latent construct, 34, 45
- Letter of intent, to foundation, 199–202
- Life Course–Social field framework
of psychological well being, 6
of social adaptation, 5
of social adaptational status, 5
- Life course development, 5
- Linguistic equivalence, in enhancing cultural relevance, 179
- LISREL, use in evaluating structural models, 37
- Logical positivism, as paradigm for quantitative methods, 145–146
- Logistical regression, use in combining information from informants, 112
- Longitudinal studies, single versus multiple publications of, 369
- Lumbar puncture (LP), 498, 503
- MacArthur (John D. and Katherine T.) Foundation, 196
- Magnetic Resonance Imaging (MRI), 288
- Management
of conflicts of interest in review process, 419–420
of constraints in informed consent, 312–314
of research budgets in, 256–257
of research with pediatric and clinical child populations, 243–260
decision making in, 253
improving data management in, 254–255
negotiating research space in, 256
negotiating time for research in, 257
record keeping in, 253
supervising research staff in, 252
of review process by editors, 411
- Mandel School of Applied Social Science (MSSAS), 541
- Manifest variable, 34–35
- March of Dimes Birth Defects Foundation, 196
- Marital variables, test of mediator in, 43
- Maternal and Child Health Branch, of Bureau of Health Care Delivery and Assistance, 216
- Measurement, 120–128; *see also* Assessment, Measurement model, Measures; Scale development
contextual approach in stress and coping, 127–133
global versus contextual approach in stress and coping, 1420–122
issues in cross-cultural research, 179
of pathology versus functional outcomes, 122–124
research directions in, 137–138
sensitivity of, 125
specificity of, 125
units of analysis, 125–126
theory based, 470–472
- Measurement model, 33; *see also* Assessment, Measures; Scale development
application of, 156–157
definition of, 45
specification of, 45
- Measures, *see also* Assessment; Measurement model; Scale development
context specific, 122
culturally relevant, 181–182
culturally sensitive, 181–182
daily diaries as example of, 123–124, 126
in clinical research, 113–114
in research proposals, 235
process oriented examples of, 123–124
scale development and modification in, 99–107
selection of, 98
self report, 98–99
bias in, 99
translations of, 99
types of, 99
- Mediation, 29–30
procedures for testing evidence of, 29–30
models of, 28–30
simple facets models of, 28–29
multiple mediator models, 30
- Mental Development Index (MDI), in Bayley Scales, 180
- Mental health services research, 464–465; *see also* Evaluating mental health services; Randomized treatment studies
conducting randomized trials in, 464
definition of, 464–465
- Meta-analysis, 395–405
appropriateness of, 399–402
criteria for evaluating, 396–404
definition of, 396
effect size in, 396, 400–403
directions for research in, 404
hypotheses in, 390–397
meaning of outcomes in, 402
practical significance of, 402–403
ruling out rival explanations in, 399–401
qualifications of, 403–404
weighing effects in, 399
- Mentorship: *see* Training researchers
in proposals to foundations, 208–209
- Millon Adolescent Personality Inventory (MAPI), 104–105
Impulse Control Scale in, 104–105
- Miami Pediatric Quality of Life Questionnaire, 289
- Moderation effect, 30, 36
definition of, 45
examples of, 41
- Mortality, role in sampling, 86
- Multicenter research: *see* Multi-institutional collaboration; Multidisciplinary research

- Multi-institutional research collaboration, 281–305; *see also* Multidisciplinary research
- Multidisciplinary research, 281–305
 benefits of, 287–292
 funding for, 289
 difficulties of, 292–300
 models of, 282–287
 role of multi-site cooperative groups in, 286–287
 professional issue in, 297–300
 authorship, 298–299
 inter-disciplinary differences in, 298
 strategies in, 293–297
 future directions in, 300–302
- Multiple informants, 107–113; *see also* Multiple measures
 combining information from, 107–113
 parent versus child, 109
 mother versus father, 108–110
- Multiple measures, 107–113; *see also* Multiple informants
 combining information from, 107–108
 complex information combining schemes in, 111–112
 use in qualitative analysis, 154–155
- Multiple regression, 26–27, 29, 32
 simple effect model of, 26–27
 test of mediation in, 29
 test of models, 36–39
- Multistage sampling, 13–16
- Multisystemic therapy (MST), 447–460; *see also* Randomized treatment studies
- Multivariate analysis of variance, 27
- Multilevel models, in analyzing individual growth curves, 56–57
 advantages of, 69–71
 stages of, 56–57
 use of in testing bias for attrition, 91–92
- National Alliance for Mentally Ill, 480
- National Bureau of Economic Research, 468
- National Cancer Institute, 216, 520
- National Commission For The Protection of Human Subjects In Biomedical and Behavioral Research, 306–307
- National Health Interview Survey on Child Health, 79
- National Institute For Child Health and Development, 216
- National Institute of Drug Administration, 216
- National Institutes of Health (NIH), 193, 196, 203, 208
 Advisory council in, 224–227
 Center For Scientific Review in, 216
 continuation grants in, 225
 deadlines for submissions to, 225
 guidelines for submissions to, 25
- National Institutes of Health (NIH) (*cont.*)
 human subjects in proposals for, 236–237
 inclusion of children in, 237
 inclusion of women and minorities in, 236–237
 multisite cooperative groups in, 286–287
 resubmission of research grants in, 225
 review in, 220–223
 study section in, 220–223
- National Institute of Mental Health (NIMH), 6, 216, 480
 Behavioral Science Research Branch of, 216
 Biometry and Epidemiology Branch of, 6
 Child and Adolescent Disorders Branch of, 216
- Division of Mental Disorders, Behavioral Research in, 8, 216
- Division of Services and Intervention Research in, 216
- Epidemiologic catchment area studies of, 8
- Multisite cooperative groups, 286–287
- National Network of Grant Makers, 193, 199
 common grant applications of, 199
- National Research Act Of 1974, 306
- National Technical Information Service, 215
- Negative case analysis, as criterion for evaluating qualitative studies, 152
- Nonprobability sampling in, 80
- Normed and non-normed fit index, 38
- Offer Self-Image Questionnaire, 104
- Ontario Child Health Study, 80
- Ordinary least squares estimates, 56
 use of in estimating individual growth curve models, 56
- Outcomes, in randomized treatment studies, 457
- Paradigms, in qualitative methods, 155
- Parameters, estimation in individual growth curves, 57–58
- Parenting Stress Index, 121
- Pediatric AIDS Clinical Trials Group (PACTG), 287
- Pediatric Oncology Group (POG), 282, 287
- Pediatrics, 350
- Pediatric settings, research in, 262–269
- Personnel, in grant application, 228–229
- Pharmacological model, 23
- Phenomenology, as qualitative method, 150
- Physical attractiveness and psychological adjustment, test of mediation model in, 42
- Physical Development Index (PDI), in Bayley Scales, 180
- Policy: *see* Integrating research and policy
- Population-based measures, 14–15
- Population-based probability sampling, 79
 advantages of, 79
 examples of, 79–80

- Population-based studies, 92
use of to enhance generalizability, 92
- Poster presentation, 343–345; *see also* Presenting scientific data
- Preliminary studies, in proposal to NIH, 233
- Presenting scientific data, 329–347
components of, 336–338
content and structure of, 336
decisions involved in, 330–332
guidelines to enhance clarity of, 338–342
pitfalls in, 342–343
review of, 332–333
tasks involved in, 330–332, 333–336
training in, 344–345
- Preventive intervention trials, 12–13
- Preventive services, 11
schema for organizing, 11
- Problem behavior theory, 31
- Process view, of change, 53–54
- Process oriented measures in chronic illness, 123
- Prognosis, impact on sample selection, 87
- Program theory, 469–470
- Progress reports
in managing research, 257
in proposal to NIH, 233
- Proposals: *see* Foundations; Research grant proposals
- Professional development: *see* Training researchers
- Professional Psychology: Research and Practice, 349
- Proximal variables, 31
in model of distress, 501
- Psychological assessment, *see also* Assessment; Measurement; Measurement model; Measures
ethical issues in research on, 314
- Psychological Bulletin, 380
- Psychological measures, *see also* Assessment; Measurement; Measurement model; Measures
issues in selecting, 98
types of, 98–99
- Psychometric issues, in cross-cultural research, 179
- Psychopharmacological models, 22
- Psychosocial Services Program, in Division of Oncology, Children's Hospital of Philadelphia, 511–526
- Publication bias, in research reviews, 389–392
- Public health perspective, 10
- Race, as a variable in research, 168–171
- Race comparative design, 171
- Rainbow Babies and Children's Hospital, Division of Behavioral Pediatrics and Psychology, 200
- Random effects
in analysis of variance, 59–65
in analysis of covariance, 60–65
in individual growth curve models, 57–60
- Randomized treatment studies, 447–460; *see also* Treatment; Treatment research
- collaboration in, 450–452, 456–457
ethical issues in, 457–460
measuring outcomes in, 457
- Multisystemic therapy(MST) as model of, 447–460
- participant characteristics in, 453–454
promoting external validity in, 453–457
promoting internal validity in, 448–453
randomized assignment in, 448–450
recruitment and retention strategies in, 454–455
- therapist selection in, 453
- treatment drop out in, 455–456
- treatment integrity in, 452–453
- treatment specification in, 451
- Random intercept
fixed slope model of, 60–65
random intercept model of, 59–65
- Raw scores, use in growth curve analysis, 71
- Reading disabled children, 52
data from in individual growth curve analysis, 52
- Recommendations
for accessing diverse populations, 174–175
for between group and within group study designs, 171–173
for designing measures in clinical research, 113–114
for enhancing data collection in pediatric settings, 265–269
for enhancing sampling, 86–92
for managing data, 114
for preparing research grants, 228–237
for research on measurement, 137–138
for sampling diverse populations, 174–175
for terminology in ethnic minority research, 168–171
for training in qualitative methods, 158–159
for training research assistants, 114
- Recruitment, *see also* Sample selection and recruitment; Sampling; Selective research participation
and eligibility in mental health services research, 474–475
and retention strategies, 454–455
establishing control of, 88–89
impact on sampling, 88–89
of ethnically diverse samples, 174–176
of underrepresented populations, 88
strategies to limit bias in, 89
- Referral bias, 84
strategies to reduce, 88
- Referral patterns, impact on sampling, 84

- Reflexivity, in evaluating qualitative studies, 153–154
- Replication, 92
as strategy to enhance generalizability, 92
- Representative sampling, in ethnically diverse populations, 173
- Research funding: *see* Foundations; Research grant proposals
- Research design, with ethnically diverse populations, 169–171
- Research grant proposals, to government agencies
- abstract in, 228
 - background and significance in, 231–232
 - biographical sketch in, 229–230
 - budget preparation in, 230
 - data analysis in, 235–236
 - data collection and procedures in, 234–235
 - deadlines for submission to NIH, 225
 - experimental design and method in, 233–237
 - helpful hints in preparing, 224–228
 - identifying options for, 213–214
 - hypotheses in, 234
 - human subjects in, 236
 - inclusion of children, 237
 - inclusion of women and minorities, 236–237
 - modifications in review process of, 224
 - personnel in, 228–229
 - pilot studies in, 213
 - preliminary studies for, 213–233
 - progress reports in, 233
 - resubmissions of, 224
 - research plan in, 231
 - resources and environment in, 231
 - securing resources for preparation of, 218–218
 - steps in preparing, 212–216
 - study sections in, 220–223
 - suggestions for preparing, 228–237
 - sampling and recruitment in, 234
 - reviewers' critique of, 237–238
 - types of, 216–217
- Mentored Scientist Development Award, 217
- Research Project Grant or Award, 217
- Program Project, 217
- Research in applied settings, 261–280; *see also* Ethics in research; Evaluating mental health services; Randomized treatment studies
- barriers to family participation in, 265
 - barriers in mental health settings, 275
 - communicating with participants in, 263–264
 - developing collaborations to enhance, 262–263, 274
 - enhancing data collection for, 265–269
 - enhancing participation in pediatric settings, 268–269
 - enhancing participation in schools, 271
- Research in applied settings (*cont.*)
- ethical issues in, 271–272
 - problems in data collection in, 264–265
 - securing resources for data collection in, 262–263
 - securing teachers cooperation for, 270
 - implementation in schools, 269–271
 - in child mental health system, 274–277
 - in child welfare settings, 272–273
 - involving practitioners, 263
- Research team, 250–252
- Response enumeration, in behavioral analytic approach, 130–131
- Response evaluation, in behavioral analytic approach, 131
- Reviewing manuscripts, 352–364, 409–423; *see also* Editing manuscripts; Writing research articles; Writing research reviews
- appeals procedure in, 420–421
 - checklist for submitting manuscripts, 357–359
 - editorial decisions in, 352–353
 - ethical issues in, 418–420
 - feedback to authors in, 351–352
 - feedback to reviewers in, 421
 - process of evaluation for, 353–356
 - reviewers' responsibilities in, 421–422
 - significance of scientific question in, 354
 - training in, 422–423
- Risk Rescue Rating Scale, 100–101
application to new population, 100–101
use with adolescent suicide attempters, 100–101
- Robert Wood Johnson Foundation, 196
- Role Play Inventory of Situations and Coping Strategies, 131–133
application to cystic fibrosis, 131–133
- Rorschach Ink Blot Test, 108
- Sample characteristics, description in published reports, 89–90
- Sample participation bias: *see* Sample selection bias
- Sample selection bias: *see* Sampling variation example of, 81
- Sample size
- in multi-institutional research, 287–288
 - in research proposals, 234
 - in testing structural equation models, 40
- Sample selection and recruitment, 80–82; *see also* Recruitment; Sampling; Sampling methods; Sampling variation
- bias in, 80–82
 - criteria in, 85–86
 - impact of attrition in, 80–82
 - impact of selective participation in, 80–82
 - impact on validity, 82
 - in research on ethnic minorities, 173–174

- Sampling, *see also* Sample selection and recruitment; Sampling methods; Sampling variation; Selective research participation
bias in, 78, 80–82
concepts of, 78–79
external population in, 78
factors affecting, 93
impact of selective participation on, 9–10
multi-stage definition of, 13–16
impact of selection of patients on, 84
sampling, 78–79
study population in, 78
target population in, 78–79
sampling error, 78–79
types of, 78–79
 sampling procedures, 79–80
 non-probability sampling, 80
 probability sampling, 70
 cluster sampling, 79
 simple random, 79
 stratified, 79
 theoretical, 79
 quota, 79
recommendations for, 87–92
- Sampling methods, *see also* Sample selection and recruitment; Sampling; Sampling variation
description of in published reports, 89–90
in qualitative research, 148–149
- Sampling variation, *see also* Sample selection and recruitment; Sampling; Selective research participation, sources of
condition related, 82–83
chronicity of condition, 82
duration of condition, 83
family demographic characteristics, 85
referral patterns, 84
setting characteristics, 84
severity of condition, 83–84
practitioners' selection of patients, 84
treatments provided to patients, 84–85
- Scale development, 99–107; *see also* Assessment; Measurement; Psychological measures
example of applying measures with new population, 99–101
example of developing short form of measure, 103
example of developing self-report measure, 103–105
example of modifying scale, 101–103
problems in developing new brief measure, 105–107
- Scale modification: *see* Scale development
- Schools, implementing research in, 269–271
- Selective research participation, *see also* Sample selection and recruitment; Sampling; Sampling variations
data analytic approaches in, 91–92
- Selective research participation (*cont.*)
 description of, 81–82
 examples of, 81–82
 in intervention studies, 82
- Selection criteria, 85–86
- Selection of measures, 98
- Setting characteristics, impact on sampling, 84–85
- Sequential analysis, 498
- Sickle Cell Disease
 Cooperative Study of (CSSCD), 282, 287–289, 293
 role of treatments in sample selection, 85
- Simple effect mediation model, 28–29
- Simple effect model, 26–27
- Simple random sampling, 79
- Single criteria construct, 35
- Site visits, role in foundation review, 206–207
- Situational analysis, in behavior analytic model, 136
- Smithsonian Scientific Information Exchange, 215
- Social adaptation, 5
- Social ecology, neighborhood effects in, 8
- Social influence models, 23
- Social Science Research Council, 533
- Society for Behavioral Medicine, 332
- Society for Developmental and Behavioral Pediatrics, 332
- Society for Research in Child Development, 332
- Southwest Regional Society of Pediatric Psychology, 332
- Socioeconomic disadvantage, as moderator of adolescent adjustment, 40–41
- Socioeconomic status, in data collection in diverse populations, 176
- Specificity
 of constructs, 33
 of measurement, 125
- Standard scores, in growth curve analysis, 71
- Stanford–Binet Intelligence Scale, application in growth curve analysis, 70–71
- Stark County Evaluation Project, 468, 471, 478
- State-Trait Anger Scale, 104
- Strategies
 for research in mental health settings, 276–277
 to enhance data collection in pediatric settings, 265
 to integrate research and practice, 425–427, 535–541
 to manage authorship in multi-institutional research, 299–300
- Stratified random sampling, 79
- Stress and coping
 contextual measurement of, 127–133
 measurement of, 121–122, 127–133
 theories of, 24
- Stroop Test, 105

- Structural equation modeling, *see also* Structural modeling
 specifying constructs in, 33–34
 use of fit index in, 37–39
- Study design, anticipating sampling problems in, 86–92
- Study population, 97; *see also* Sampling; Sampling variation
- Study section, in research grant review, 220–223
- Suicide Intent Scale (SIS), 100–101
 use with adolescent suicide attempters, 100–101
- Target population, 78
- Task equivalence, 177, 179–182
 in research with ethnic minority children, 167
- Terminology, in research with ethnic minority children, 167
- Theoretical frameworks, 21–50; *see also* Theory in child health research, 21–25
 in developmental epidemiology, 3–17
- Theoretical sampling, 79, 154
 as criteria for evaluating qualitative studies, 154
- Theoretical significance, of research reviews, 354
- Theories of implementation, 414
- Theory, *see also* Theoretical frameworks
 types of, 25
 buffering, 11
 diathesis stress, 23
 family systems,
 grounded, 150
 social influence, 23
 stress and coping, 24
- Theory-based measurement, 470–472
- Theory-driven research, 12–50
- Time-controlled methods, role in managing attrition, 92
- Trail Making Test, 105
- Training researchers
 Boulder model in, 112
 in assessment, 137
 in case study method, 443–444
 in conducting research in clinical settings, 278–279
 in epidemiology, 15–17
 in ethical conduct of research, 320–321
 in ethnic diversity, 182–184
 in health behavior research with minorities, 184
 in interdisciplinary research, 300–301
 in managing research, 257–260
 in manuscript editing, 422–423
 in manuscript review, 422–423
 in preparing research proposals for foundations, 208–209
- Training researchers (*cont.*)
 in preparing research proposals for government agencies, 238–240
 mentorship in, 239–246
 grant writing courses in, 239
 in public health perspectives, 15–17
 in scientific presentations, 344–345
 in scientific writing, 369–372
 role of writers' workshop in, 369–372
 in qualitative methods, 157–159
 in writing review articles, 392–393
 role of writers' workshop in, 209
 role of Mentored Research Scientist Development Award in, 297
 role of mentorship in, 205–209
- Transferability, as criteria for evaluating qualitative studies, 153–154
- Treatment 451–452; *see also* Randomized treatment studies
 dropout in studies of, 455–456
- Treatment research, 492–494; *see also* Evaluating mental health services, Randomized treatment studies
- Treatment services, schema for organizing, 11
- Triangulation, use of in qualitative methods, 154–155
- Trustworthiness, as criterion for evaluating qualitative studies, 152
- Tucker Lewis Index, 38
- Unconditional model, in individual growth curve analysis, 57–66
- Underrepresented populations, recruitment of, 88–89
- United States Government Manual, use of in identifying options for research grants, 213
- U.S. Government Printing Office, 213
- Urban Child Health Initiative, 474
- Ways of Coping Check List, 121
- Wechsler Intelligence Scale For Children, application in growth curve analysis, 70–71
- Weighted least squares, in analyzing individual growth curves, 57
- William T. Grant Foundation, 182
- Within group studies, of ethnic minority populations, 170–173
- Writers' workshop seminar
 in preparing proposals for foundations, 208
 in preparing grants to government agencies, 239
 in training for scientific writing, 369–371
- Writing grants, courses in, 239
- Writing research articles, 340–361; *see also*
 Editing manuscripts; Reviewing manuscripts; Writers' workshop seminar;
 Writing research reviews
 audience for, 349–350
 focus of, 349

- Writing research articles (*cont.*)
guidelines in, 364–365
managing difficult issues in, 367–369
methods in, 359–361
options in, 350–357
procedures for informed consent in, 362
procedures in, 361–363
research findings in, 356–364
results in, 362–363
reviews of, 351–356
revisions of, 365–369
use of feedback checklist for authors in, 357–360
use of formatting checklist for authors in, 360–361
- Writing research reviews, 377–394; *see also*
Editing manuscripts, Reviewing manuscripts
accessing studies for, 382–383
dilemmas in, 387–389
focus of, 379–380
- Writing research reviews (*cont.*)
generalizability of findings in, 392
hypotheses in, 384–385
introduction in, 381
implications of, 386–387
inclusionary criteria in, 383–384
exclusionary criteria in, 383–384
methods in, 382–384
methodological limitations of, 369–372
presentation of findings in, 385–387
publication bias in, 389–392
publication outlets for, 380
purpose of, 381–382
sample format for, 390–391
structure of, 380
tasks in preparation of, 376–378
theory in, 385
training in, 392–393
types of, 378–379
- Youth Health Risk Behavior Inventory (YHBI), 172