National Institute on Drug Abuse

RESEARCH

MONOGRAPH SERIES

Drug Abuse Prevention Intervention Research: Methodological Issues

107

U.S. Department of Health and Human Services • Public Health Service • Alcohol, Drug Abuse, and Mental Health Administration

Drug Abuse Prevention Intervention Research: Methodological Issues

Editors:

Carl G. Leukefeld, D.S.W.
William J. Bukoski, Ph.D.

NIDA Research Monograph 107
1991

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Alcohol, Drug Abuse, and Mental Health Administration

National Institute on Drug Abuse
5600 Fishers Lane
Rockville, MD 20857

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, DC 20402

ACKNOWLEDGMENT

This monograph is based on the papers and discussion from a technical review on "Drug Abuse Prevention Intervention Research: Methodological Issues," held on May 22-23, 1989, in Rockville, MD. The review meeting was sponsored by the National Institute on Drug Abuse.

NIDA Research Monographs are indexed in the *Index Medicus*. They are selectively included in the coverage of *American Statistics Index, BioSciences Information Service, Chemical Abstracts, Current Contents, Psychological Abstracts,* and *Psychopharmacology Abstracts.*

DHHS publication number (ADM)91-1761
Printed 1991

Contents

140136

An Introduction to Drug Abuse Prevention Intervention Research: Methodological Issues

Carl G. Leukefeld and William J. Bukoski

INTRODUCTION

With the renewed emphasis on drug abuse prevention, questions now are being asked about the effectiveness of those prevention interventions. Responses to these questions clearly suggest that drug abuse prevention interventions have been inconsistent in changing drug abuse and related behaviors. It has been suggested that more is known about what does not work than about what works in preventing drug abuse (Berberian et al. 1976; Goodstadt 1974; Schaps et al. 1981). Nevertheless, much is known about drug abuse prevention, and there are promising approaches (Donohew et al., in press; Glynn et al. 1983).

Most questions related to the effectiveness of drug abuse prevention interventions center on research design and methodology and on the differences as well as the inconsistencies among study findings. An example is the choice of outcome measures (e.g., no drug use as contrasted with occasional drug use), which significantly affects a study's findings and consequently a study's importance. Additional methodological issues are important to the prevention practitioner and the researcher; these issues are the basis of this volume. Moreover, there are questions and discussion about the limitations of drug abuse prevention evaluation research and prevention evaluations in general (Biglan and Ary 1985; Leukefeld, in press) as well as recommendations that new research methodologies be developed to better understand prevention interventions. Some suggest that drug abuse prevention researchers talk more frequently with prevention practitioners about their expectations, opinions, and anecdotal experiences related to prevention program effectiveness.

Drug abuse prevention has been controversial; consequently, drug abuse prevention research is a part of that controversy. Swisher (1979) identified the following controversial issues: evidence that prevention makes a difference;

difficulty in agreeing on how to demonstrate the effectiveness of prevention strategies; confusion regarding the differences among treatment, intervention, and prevention efforts; and concern about the purpose of prevention—ultimate use or nonuse by the target population. Silverman (in press) confronts this issue by suggesting that prevention research is evolving and will continue to develop as a direct consequence of more complex theoretical and conceptual thinking, more valid and reliable measures of drug-related problems, better understanding of individual risk factors, better identification of individuals and groups at high risk, better research design and long-term followup studies, and better integration across various settings—family, school, community, religious, and criminal justice.

With this brief background and the overriding goal of providing greater clarity to research findings, the purpose of this monograph is to examine the state of the art of drug abuse prevention research methodology, to develop recommendations for refining current methodological approaches, and to develop an agenda for future research applications. Authors were asked to emphasize instrumentation, control/comparison groups, intervention specificity, clarification of outcome variables, replication issues, and measurement of long-term effects. Although this monograph does not review drug abuse prevention effectiveness research, chapter authors were asked to present data and research findings as examples of methodological issues.

DRUG ABUSE PREVENTION RESEARCH

Defining prevention is a first step in exploring drug abuse prevention methodology. Bukoski (in press) identifies three approaches or perspectives for drug abuse prevention. First, the public health model incorporates the concepts of primary, secondary, and tertiary prevention (Last 1980). Primary prevention is directed to preventing the onset of disease, including decreasing the incidence, new start, or onset. Many drug abuse prevention activities can be placed in this category of primary prevention. A common criticism of this conceptualization of prevention is the overlap between categories as well as the fact that all prevention, treatment, and rehabilitation services could be categorized within this definition of prevention.

A second prevention conceptualization, the communicable disease model, focuses on the host, agent, and environment (Wilner et al. 1978). (The agent is the germ, virus, or other cause of a disease.) The host relates to the human susceptibility or resistance to disease and can be influenced by many hereditary and lifestyle factors. Environment refers to social or physical factors that may contribute to the initiation and spread of a disease.

A third conceptualization is the risk factor model, which is directed to identifying psychological, social, and biologic factors related to the emergence of a health problem (Arnold et al. 1981). Risk factors have been used extensively to depict increased risks for drug abuse using correlation research and other research findings. From one point of view these three ways of thinking provide some clarity about prevention, but from another point of view they also generate confusion because they are not completely compatible.

Defining prevention research has not always been clear. Bukoski (1980) adds clarity with a description of a drug abuse prevention research evaluation model that includes three levels of evaluation: (1) process evaluation, which focuses on assessing the service operation of a prevention program and includes descriptions of the program's prevention services, use of resources, and costs; (2) outcome evaluation, which is used to determine if a prevention program's objectives were met by applying comparative evaluation designs; and (3) impact evaluation, which is used to assess macroindicators of drug abuse at the community level. The outcome evaluation design is probably the most frequently used design for evaluating drug abuse prevention. This controlled and comparative design of two or more groups is reviewed by Snow and Tebes in this volume.

THE IMPORTANCE OF RESEARCH METHODOLOGY

This volume is organized into five areas: introduction, documenting the prevention intervention, measuring the efficacy of prevention interventions, assessing effectiveness, and consensus development. Clayton and Cattarello provide an overview of drug abuse prevention research and discuss several methodological issues from their current research, which focus on assessing the Kentucky drug abuse resistance education prevention program. Johnston reviews drug trends among senior high school students and presents an overview of prevention impact in the United States.

In the section on documenting prevention intervention, Flay and Petraitas provide a rationale and argue for the need to base prevention intervention programs on a strong theoretical foundation. History suggests that many prevention interventions, focused on drug abuse and other health promotion areas, have skipped this important part of documenting the intervention. Gilchrist discusses the parameters for defining a prevention intervention and for delineating the intervention target audience. Defining the prevention intervention as the intervening variable is not only essential for program replication but also is important for process evaluation and for training intervention staff. Specifying the target audience is also important in understanding the limitations of the intervention and the anticipated outcomes.

Pentz and Trebow present the final chapter in this section by reporting on issues related to program implementation. The implementation of drug abuse prevention program interventions is influenced by many environmental variables, which along with other factors make drug abuse prevention intervention research interesting and dynamic.

The third section focuses on measuring the efficacy of prevention interventions. Snow and Tebes provide a review of experimental and quasi-experimental research designs that can be used in prevention intervention studies. Validity, basic threats to validity, and tradeoffs are examined. Bentler presents an overview of modeling and measurement issues related to measuring the effects of the prevention intervention. Using statistical controls is suggested when research designs break down. Dwyer and MacKinnon discuss outcome measures used in drug abuse prevention intervention research. Areas for consideration include potential variable type, issues related to validity, categorical variables, and pretest measures.

In the section on assessing the effectiveness of drug abuse prevention interventions, Hawkins and colleagues examine the long-term effects of drug abuse prevention interventions and issues related to replication. Prevention intervention effects are temporally limited, and approaches need to be further refined to enhance impact (e.g., booster sessions). Biglan and coworkers review issues related to controlling and examining attrition. Study dropouts need to be better understood; generally, those with higher rates of problem behaviors, including drug abuse, tend to leave prevention studies. Likewise, the effects of those entering a study (drop-ins) need to be assessed. Forman and Linney present approaches for validating drug abuse self-reports, which include physical/chemical tests, behavioral observations, and peer ratings. Validating outcome measure is an issue that should be considered in all drug abuse prevention intervention research. Schinke and Orlandi present stages for transferring technologies that focus on drug abuse prevention. Technology transfer stages incorporate a range of possibilities from basic research to adaptation to obsolescence.

REFERENCES

Arnold, C.; Kuller, L.; and Greenlick, M., eds. *Advances in Disease Prevention.* Vol. 1. New York: Springer Publishing Company, 1981.

Berberian, R.M.; Gross, C.; Lovejoy, J.; and Paparella, C. The effectiveness of drug education programs: A critical review. *Health Educ Monog* 4:377-398, 1976.

Biglan, A., and Ary, D.V. Methodological issues in research on smoking prevention. In: Bell, C.S., and Battjes, R., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 170-195.

Bukoski, W. Where drug education stands now. In: Edwards, G., and Arif, A., eds. *Drug Problems in the Sociocultural Context.* Geneva, Switzerland: World Health Organization, 1980.

Bukoski, W. A definition of drug abuse prevention research. In: Donohew, L.; Sypher, H.; and Bukoski, W., eds. *Persuasive Communication and Drug Abuse Prevention.* Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Donohew, L.; Sypher, H.; and Bukoski, W., eds. *Persuasive Communication and Drug Abuse Prevention.* Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Glynn, T.J.; Leukefeld, C.G.; and Ludford, J.P., eds. *Preventing Adolescent Drug Abuse: Intervention Strategies.* National Institute on Drug Abuse Research Monograph 47. DHHS Pub. No. (ADM)83-1282. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983.

Goodstadt, M.S. Myths and methodology in drug education: A critical review of research evidence. In: Goodstadt, M.S., ed. *Research on Methods and Programs of Drug Education.* Toronto, Ontario: Addictions Research Foundation of Ontario, 1974.

Last, J. *Public Health and Preventive Medicine.* 11th ed. New York: Appleton-Century-Croft, 1980.

Leukefeld, C.G. The role of the National Institute on Drug Abuse in drug abuse prevention. In: Donohew, L.; Sypher, H; and Bukoski, W., eds. *Persuasive Communication and Drug Abuse Prevention.* Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Schaps, E.; DiBartolo, R.; Moskowitz, J.; Palley, C.; and Churgin, S. Primary prevention evaluation research: A review of 127 impact studies. *J Drug Issues* 11:17-43, 1981.

Silverman, M.M. *Prevention Research: Impediments, Barriers, and Inadequacies. Proceedings of the First National Conference on Prevention Findings.* Washington, DC: U.S. Govt. Print. Off., in press.

Swisher, J.D. Prevention issues. In: Dupont, R.L.; Goldstein, A.; and O'Donnell, J., eds. *Handbook on Drug Abuse.* Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1979.

Wilner, D.; Walkley, R.; and O'Neil, E. *Introduction to Public Health.* 7th ed. New York: MacMillan, 1978.

AUTHORS

Carl G. Leukefeld, D.S.W.
Director
Center for Drug and Alcohol Research
Department of Psychiatry
University of Kentucky College of Medicine
Annex 2, Room 206
Lexington, KY 40536-0080

William J. Bukoski, Ph.D.
Program Director
Intervention Research
Prevention Research Branch
Division of Epidemiology and Prevention Research
National Institute on Drug Abuse
Rockwall II, Suite 615
5600 Fishers Lane
Rockville, MD 20857

140137

A Framework for Drug Abuse Prevention Research

William J. Bukoski

INTRODUCTION

Research interest in the efficacy and effectiveness of drug abuse prevention programs has increased dramatically over the past 10 years, resulting in closer scrutiny of the quality and appropriateness of research methods, tests and measures, as well as data analysis procedures used for program evaluations. Efficacy refers to determining treatment effects resulting from an experimental assessment that has internal validity. Effectiveness research assesses the generalizability of the intervention when implemented under real-world conditions. Concern about conducting quality prevention research has been prompted by increased public recognition of the drug abuse problem, growing pressure to take effective action, and a substantial increase (since 1986) in Federal support for drug-related law enforcement, treatment, and prevention to fight the "war on drugs" (Office of National Drug Control Policy 1990).

Health policymakers, legislators, community leaders, and concerned citizens are seeking scientifically based answers to their questions about which prevention strategies work best, for whom, in which situations, and for what length of time. Immediate and conclusive answers to these critical questions, however, are still emerging from the research. In part, research that focuses on the efficacy of drug prevention programs has been hampered by scarce resources, including an insufficient number of prevention research scientists to conduct needed studies and conceptual ambiguities about the nature of drug abuse prevention programing and prevention research. Although the first issue has practical solutions that involve increased resources, the other area requires discussion of the interrelation between prevention concepts and the design of technically sound prevention research. Pertinent to this discussion are two basic questions: What is the theoretical basis for designing drug abuse prevention interventions? and What are the most appropriate research methods to assess the efficacy and effectiveness of prevention programs? In response, this chapter proposes a model for prevention programing based on etiologic

research and a framework for drug abuse prevention research that incorporates process, outcome, and impact prevention research methodologies.

DEFINING DRUG ABUSE PREVENTION

Progress in drug abuse prevention research has been hampered in part by the need to develop a clear definition of "prevention" that has the consensus of practitioners, researchers, and policymakers (Bukoski 1990). In the past, questions have been raised about the objectives of drug abuse prevention programs. For example, should prevention programs be designed to prevent the initiation of a drug by a nondrug user, or should a prevention activity also interrupt an individual's progression from the use of alcohol and/or cigarettes to marijuana, amphetamines, cocaine, and other drugs of abuse? Questions have been raised about the appropriate target audiences for prevention programs. Should drug abuse prevention programs focus primarily on the individual, or would multiple audiences be appropriate targets of prevention, including parents, other family members, peer groups, or larger social environments such as schools, communities, and the workplace? Finally, the substantive content of drug prevention has been widely debated over the past 20 years with different segments of the field advocating approaches that ranged from scare tactics; through effective education designed to strengthen self-concept, social skills development, and community organization and action; to deterrence via social control measures and punitive consequences (Glynn et al. 1983; Bukoski 1986).

As a result, the field of drug abuse prevention has not clearly delineated (1) the theoretical basis for programs; (2) specific, measurable, and predicted program outcomes; or (3) probable impact of programs on drug use incidence and prevalence when measured within a program's service area or the community at large. Scientifically based discussion of these issues serves as the benchmark for the design of sensitive and authoritative evaluation research projects. Without this information, researchers and program officials who want to evaluate a program's effect may have to make important decisions concerning research hypotheses, dependent and independent variables, test instruments, data collection protocols, and data analytic procedures without clear guidance from the scientific literature.

ISSUES IN PREVENTION RESEARCH

A second issue in prevention research has been implementing sound research methodologies (Schaps et al. 1981). Many early prevention evaluations were poorly designed and did not measure or were insensitive to assessing the effects of drug prevention programs. The prevention research literature reveals

that early research designs frequently incorporated off-the-shelf instruments that may have been inappropriate to the prevention program's actual objectives, did not distinguish between process and outcome effects, and used experimental research designs that were incapable of producing meaningful and interpretable program outcome data (Leukefeld and Moskowitz 1983; Goodstadt 1986; General Accounting Office 1988).

In part, the traditional solution of using controlled experimental designs may have been inappropriate for evaluating drug abuse prevention programs. As a result, many early drug abuse prevention evaluations suffered by force-fitting an experimental research paradigm—with marginal degrees of precision—on a prevention program activity that often was developed poorly consistent with a theory of predicted effect. Frequently, the research design did not include the measurement and analysis of theoretically relevant process or mediating variables, did not assess the quality and quantity of program implementation, and did not examine critical relationships such as potential subject by treatment interactions, effects of differential attrition, and changes in the normative climate that may have affected possible program outcomes. Many early controlled research prevention studies focused on one rather than multiple outcomes that were expected to change (e.g., positive changes in self-concept, decisionmaking and communication skills, drug knowledge, and lower levels of the incidence and/or prevalence of drug use and abuse by program recipients). Most important, prevention studies frequently did not use available research findings to assess program theory or to guide the design of future prevention interventions. Viewed from this perspective, it is possible that the apparent failure by drug abuse prevention programs in producing consistent and enduring drug use effects could be related to a variety of factors, including theory, implementation, poor research designs, or a combination of these influences (Schaps et al. 1984).

CURRENT DEVELOPMENTS IN DRUG ABUSE PREVENTION PROGRAMING

Over the past 5 years, more time, attention, and research have focused on applying etiologic research to designing and testing theory-based drug abuse prevention interventions. This scientific literature suggests that a single, "silver bullet" preventive solution was not supported by the research. Rather, research indicates that drug use and abuse have multiple causes and correlates. Although current drug abuse prevention models more clearly define specific risk factors to drug use onset and progression that would be addressed by a program, a comprehensive theory for drug abuse prevention is needed (Bukoski, in press).

Role of Etiologic Research

Designing prevention interventions requires knowledge of etiologic risk factors for drug use onset, progression, and abuse. A risk factor approach focuses on the identification of those psychological, social, biologic, behavioral, and environmental factors that appear to be correlated with the emergence of a health problem. The term "risk factor" was used by Stamler in 1958 (Arnold et al. 1981) and was initially applied to cardiovascular disease prevention (Stamler 1978, 1979). This literature suggests that less exposure to salient risk factors may serve to protect or inoculate youth against the subsequent use and abuse of drugs (Simons et al. 1988). The risk factor approach quickly became a valuable and popular approach in public health, including drug abuse prevention.

Over the past 20 years, many research studies have examined factors related to drug use and abuse. For example, research by Gorsuch and Butler (1976), Bry (1983), Hawkins and colleagues (1986), Newcomb (1988), Cloninger (1988), Schuckit (1987), and Pickens and Svikis (1988) provides a theoretical basis and empirical structure for the scientific understanding of the drug abuse causes and guidance in designing and testing preventive interventions. Etiologic research studies support the view that several pathways to drug use and abuse occur and that there is not one simple reason why youth may be vulnerable to drug use and abuse (Jones and Battjes 1985).

Etiologic research suggests that prevention research should address risk factors across at least four clusters: individual, family, peer group, and community (school, workplace, and local neighborhood) (figure 1). Although simplistic in structure, this model suggests that many specific risk factors may play a role in drug use onset and progression and that these factors may be dynamically related within and across categories.

Individual drug abuse risk factors include early drug use; nonconventionality; inadequate social bonding; deviant behavior; adult, parental, or older sibling role models who use drugs; novelty- or sensation-seeking; personality factors such as early signs of aggressive and or noncompliant behavior; low religiosity; low academic achievement; psychological distress or depression; and low self-efficacy or self-acceptance (Newcomb et al. 1986). Survey research indicates that an important risk factor is a youth's misperception of the harmful consequences and social disapproval of drug use (Johnston 1985). Bachman and colleagues (1988) report that the downward trend in marijuana use from 1978 to 1986 was best explained by positive changes in these two variables.

Drug Abuse Risk Factors

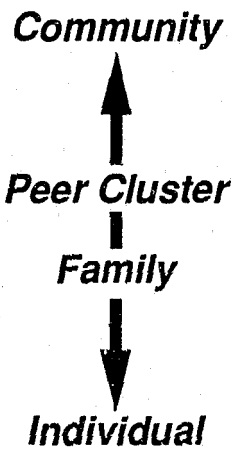

FIGURE 1. *Etiologic risk factors*

Family factors include a history of alcoholism and antisocial behavior, parental and older sibling drug-using role models, ineffective parenting practices, and lack of mutual parent-child attachment and warmth (National Institute on Drug Abuse 1987).

Peer factors include peer drug use (Newcomb et al. 1986), peer cluster influence through social interaction (Oetting and Beauvais 1987), and peer social pressure (Brown et al. 1989).

Community factors include availability of drugs and alcohol (Rush et al. 1986); drinking and driving laws and their reinforcement/enforcement, alcohol price, and minimum drinking age laws (National Institute on Alcohol Abuse and Alcoholism 1987); social/cultural norms and mores relevant to use; social stress (Linsky and Straus 1986); and lack of economic mobility and social supports and poverty (Auslander 1988).

Researchers have developed complex prevention theories that further articulate subcategories of etiologic factors. For example, Kumpfer (1987) proposes a biopsychosocial vulnerability model that suggests that genetic and biologic factors involving parents' interaction with prenatal and early childhood development play an important early role in shaping adolescence (figure 2).

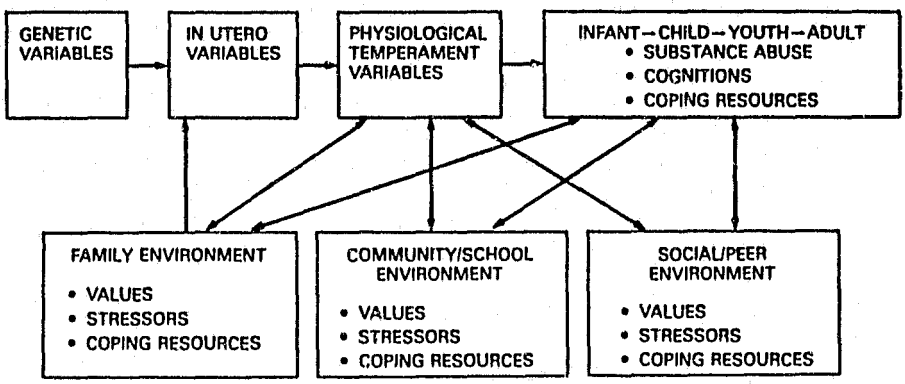

FIGURE 2. *A biopsychosocial vulnerability model*

Prevention theory also is guided by psychopharmacology studies. The relationship of this knowledge base to other psychosocial research is clearly depicted by the public health model of contagious diseases (figure 3) that shows a reciprocal relationship between host (individual), environment (biologic, social, and physical), and agent (drugs). At least four implications for the design of preventive interventions are suggested by this model (Arnold et al. 1981): (1) increase individual resistance to the agent (e.g., peer resistance training); (2) protect individuals from the agent (e.g., drug education, abstinence model); (3) isolate the agent from the host (e.g., establish drug-free school/ community zones); and (4) modify the agent to reduce risk of harm (e.g., lower or eliminate alcohol content of beverage, encourage use of filter-tipped cigarettes).

Comprehensive Drug Abuse Prevention

Because drug abuse is a progressive and chronic relapsing disorder, as well as a health problem with multiple pathways, it may be necessary to target varied preventive strategies at different stages of the emerging problem. It also appears that simultaneous focus should be placed on preventive strategies, including the individual, the family, the peer group, and the community as well as schools, workplaces, and local neighborhoods (Bukoski, in press).

Comprehensive drug abuse prevention offers a combination of strategies consistent with individual needs and developmental levels, while sequencing these interventions consistently with each appropriate stage of drug use behavior (figure 4). This approach recognizes that drug abuse encompasses a

COMMUNICABLE DISEASE MODEL

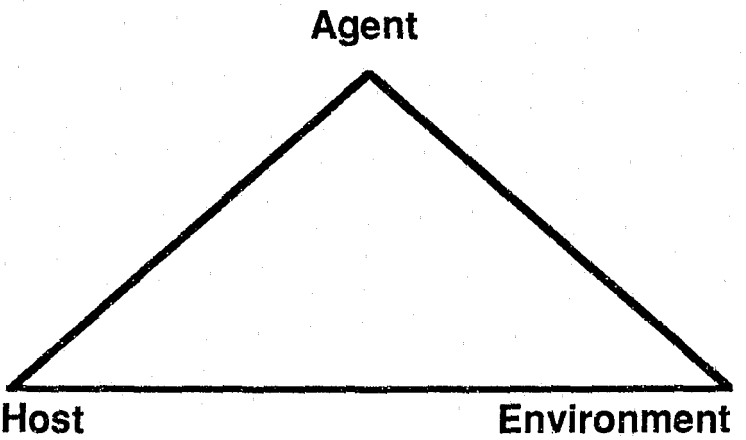

FIGURE 3. *The public health model*

spectrum of behaviors from nonuse to dependency and includes a comparable range of theoretically based prevention strategies along this continuum of drug use.

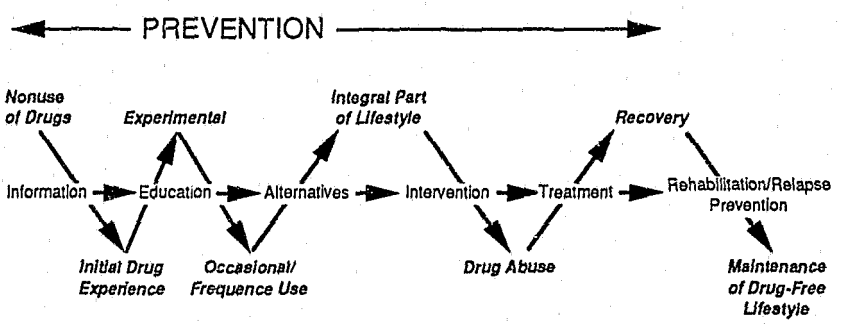

FIGURE 4. *A mode of comprehensive prevention*

Some view primary prevention as the transition from nonuse to initial or first use. An alternative position is that prevention programs need to focus on youth and young adults who are moving through stages of drug experimentation (i.e.,

occasional/frequent use, integrating drug use as part of their lifestyle to the point of medically diagnosed drug abuse). At the point of medical diagnosis of drug abuse, prevention as a health program concept would end and drug treatment would begin.

Comprehensive drug abuse prevention involves multiple program components appropriate to individual, family, peer group, and community. Four types of prevention programs can be described.

Information programs describe the harmful physical and psychological consequences of drug consumption. Information programs include media campaigns, drug education lectures, films, pamphlets, flyers, bumper stickers, and media coverage of drug-related events (Shoemaker 1989).

Education programs are designed to remediate deficiencies in social and psychological skills, improve interpersonal communication, promote self-understanding and acceptance, and master refusal training to counter a variety of social influences to use drugs. Research indicates that academic success and achievement motivation may serve as a protective factor against drug use and abuse (Hawkins et al. 1988; Brook et al. 1986) and that comprehensive drug prevention programs include activities to improve educational attainment through techniques such as mastery learning (Dolan et al., in press).

Alternative programs provide opportunities to individuals who may be at risk of drug use because of a need for excitement or sensation and a socially acceptable and authentic way of offsetting boredom or dissatisfaction with one's life. Positive alternatives to drug use have included Outward Bound or wilderness experiences, cooperative community service or restoration projects, skydiving, and volunteering time and talents to help another person (Cook et al. 1984; Cook 1985; Tobler 1986).

Intervention programs are appropriate for high-risk individuals who need special assistance to recognize the signs and symptoms of initial drug and alcohol dependency and corrective or rehabilitative actions that may take the form of crisis intervention or drug hotlines, peer counseling, peer leadership programs, parent peer groups, or psychological counseling at the individual or family level (Tobler 1986; Morehouse 1979; Myrick and Erney 1979; National Institute on Drug Abuse 1981; Manatt 1983). Interventions also can focus on the community. Examples include social policies to create a drug-free workplace and implementation of appropriate physiological testing, enactment of drug-free school/community zones legislation, and drug-related law enforcement operations within the community (National Institute on Drug Abuse 1989; New Jersey Department of Public Safety 1988). Recent research indicates that with

multiple pathways to drug use and abuse, the most effective prevention approach would incorporate multiple component programs to address several salient risk factors within the same program (Pentz et al. 1989).

A MODEL OF COMPREHENSIVE DRUG ABUSE PREVENTION

In 1984 the National Institute on Drug Abuse (NIDA) initiated a drug and alcohol research program designed to promote multiple component, multiple level, comprehensive drug abuse prevention program research. As a result of that grant program, NIDA awarded a 5-year drug abuse prevention research grant to the University of Southern California to assess the efficacy of a comprehensive approach to drug abuse prevention at the community level. The project, titled the Midwestern Prevention Project, is located and involves all the schools and communities in the Kansas City, MO, and Indianapolis standard metropolitan statistical areas. Joint funding for program delivery is provided by the Lilly Foundation, Kaufman Foundation, and Marion Laboratories. The research in Kansas City is a quasi-experimental design with nearly 50 public middle and junior high schools matched on demographic characteristics and assigned to either a treatment or control condition. In Indianapolis, the research randomly assigned nearly 60 schools to either treatment or control conditions.

Five interventions are being tested sequentially in each site, with one new intervention added each year. The first intervention is mass media, which is used each year of the project to heighten community awareness to the drug abuse problem and to introduce the new intervention being implemented by the program. During the first year, the project introduced a school-based peer resistance program called STAR (Students Taught Awareness and Resistance), which consists of 10 classroom and homework sessions focused on the psychosocial consequences of drug use; correction of normative myths concerning the prevalence of drug use by teens; social resistance training to offset and counter adult, media, peer, and community influences to use drugs; assertiveness and problemsolving training; and a statement of public commitment to avoid using drugs. Methods for delivery include the use of modeling and rehearsal (role-playing) of resistance skills; corrective feedback on skill development from the instructor and peer group; homework assignments that involve parent-child discussions concerning the problems resulting from the consumption of tobacco products, alcohol, and other drugs; problemsolving of difficult social pressure situations; and peer leader facilitation of the teacher-implemented program.

During the first year, mass media events included a total of 16 television, 10 radio, and 30 print media ads that described the school component and discussed the drug problem in the community. The remaining three

interventions include parent/family drug education and organization, community organization activities through creation of a drug council or task force, and efforts to change health-related policies in the community. Research results from the first project year in Kansas City indicate that sixth- and seventh-grade program participants report significantly lower levels of alcohol, cigarette, and marijuana use at 1-year followup than students in the comparison condition controlling for race, grade, socioeconomic status, and urbanicity (17 percent vs. 24 percent for cigarette smoking; 11 percent vs. 16 percent for alcohol use; and 7 percent vs. 10 percent for marijuana use in the past month) (Pentz et al. 1989). Preliminary, unpublished followup data indicate that these initial differences are maintained over the following 4 years and that the positive effects of the program may extend to the use of cocaine as well (Pentz, in press).

A DRUG ABUSE PREVENTION RESEARCH MODEL

If drug abuse prevention requires a comprehensive approach then it follows that prevention research requires a more encompassing research methodology to assess the efficacy and effectiveness of comprehensive prevention programing (French and Kaufman 1984). The prevention metaresearch framework proposed in this chapter integrates etiologic, intervention, and epidemiological research by linking theoretical studies on the causes of and mediating faciors relevant to drug use and abuse (process research), controlled efficacy studies of theory-based prevention interventions (outcome research), and community epidemiology (impact research to assess in the aggregate preventive effects over time) (figure 5).

Metaresearch includes process, outcome, and impact methods and attempts to develop answers to three basic questions: (1) What was the theory, social/ cultural context, program content, intensity, and quality of the preventive intervention being tested, and what was its level of fidelity to its underlying theory? (2) What was the efficacy of the intervention relevant to predicted changes in drug-related knowledge, attitudes, and behaviors? and (3) What was the generalizability of the research to the larger population, and to what extent was the intervention effective when administered under realistic real-world conditions typical of schools, the workplace, medical/health clinics, and communities?

Process Research—Theory Testing and Documenting the Intervention

Process research delineates the theoretical basis for the intervention and describes in comprehensive fashion the program activities planned and implemented to achieve the predicted changes in drug-related attitudes,

LEVEL OF EVALUATION	PROCESS →	OUTCOME →	IMPACT
Focus of Evaluation	Prevention Program Effects		Aggregate or Cumulative Effects at the Community Level
Potential Indicators Of Effectiveness	Description of Target Audience/Recipients of Service Prevention Services Delivered Staff Activities Planned/ Performed Financing Resources Utilized	Changes In Drug-Related: - Perceptions - Attitudes - Knowledge - Actions: Drug Use Truancy School Achievement Involvement in Community Activities	Changes In: - Prevalence and Incidence of Drug Use - Drug-Related Mortality/ Morbidity - Institutional Policy/ Programs - Youth/Parent Involvement In Community - Accident Rates
Potential Prevention Evaluative Approaches	Examples: - Management Analysis - Ethnography - Qualitative Methods - Cost Accountability Model - Quality Assurance Assessment - Quality Circles	Examples: Experimental Paradigms Quasi-Experimental Designs Ipsative Designs (e.g., Goal Attainment Scaling)	Examples: Epidemiologic Studies Incidence and Prevalence Studies Drug-Related School Surveys Cost-Benefit Analysis

FIGURE 5. *A prevention research model*

knowledge, and behaviors. In experimental research terms, process research is the equivalent of measuring the "independent variable" (Schuerman 1983).

Verification of the independent variable is important for three reasons (Shaver 1983). First, researchers should not draw conclusions concerning the efficacy of an intervention without first confirming that the independent variable was implemented. Second, accurate replication of research requires a thorough knowledge of the experimental conditions and setting. Third, synthesis of research findings is facilitated because data drawn from research studies that include accurate and well-defined interventions can be compared more accurately and comprehensively. Most importantly, process research provides data necessary to test the validity of theory and to evaluate the integrity and merit of program implementation. Process research attempts to answer several salient questions.

- What was the proposed theoretical intervention basis, and what was the level of fidelity to theory?

- What effect did program exposure have on theory-based mediating variables?

- What was the content, intensity, and intervention quality? Was it implemented as planned? Did the intervention provider modify elements of the program to meet his or her unique styles or the subjects' perceived needs? Did the subjects attend all program sessions? Did the subjects actively participate in all sessions in a uniform fashion? What was the program content delivered in the test site? Was the content varied either in terms of sequencing or extent of coverage?

- Was the intervention relevant and appropriate to the target audience from a theoretical, developmental, social, and cultural perspective?

- Was the program fiscally accountable and efficiently operated?

Unlike laboratory studies in the physical sciences where the quantity and quality of the independent variable is well controlled, a prevention intervention trial usually is conducted in field settings (e.g., schools, communities, and the workplace). Process research helps to identify the source, magnitude, and potential confounding effects of social and contextual influences on the intervention.

Process research captures in still-frame fashion the operational features and dynamics of a prevention program through qualitative and quantitative analysis.

Process research also identifies and measures unanticipated or unplanned events and assesses their effects on planned program operations. Process research provides a measure of the integrity of the preventive intervention and describes it in terms of its theory, objectives, and implementation, making it possible to determine why a program's intended outcomes were or were not achieved. In essence, the purpose of process research is to delineate an intervention's reality.

Process evaluation has two other valuable uses. First, process research results could be used to accurately replicate the intervention research in a different setting or with a different population. Second, if the program were judged successful and a decision were made to include the intervention prevention services, process research findings would help identify which program components were essential and should be retained in the broad-scale implementation and which program elements appeared incidental to achieving positive effects (Jason et al. 1986). If the program proved to be ineffective, process research also would help to determine whether the results were due to program theory failure, inadequate implementation, or a combination of both.

Other chapters in this volume (Flay and Petraitis, Gilchrist, Hawkins and colleagues, and Pentz and Trebow) clearly elucidate the importance of appropriate process research methods.

Outcome Research—Measuring the Efficacy of Preventive Interventions

The purpose of outcome research is to assess the intervention's efficacy to effect positive changes in dependent variables. Variables of interest may include drug-related knowledge, attitudes, and beliefs; perceptions of harmful consequences and social disapproval; self-reported or actual levels of drug use; and drug-related behaviors such as truancy, school underachievement, or delinquent acts. Design and implementation of scientifically sound outcome research are primarily concerned with maximizing internal validity, which refers to the capacity of the outcome research to directly link changes in relevant dependent measures to participation in the experimental intervention, rather than to unmeasured variables or extraneous events (French and Kaufman 1984; Hawkins and Nederhood 1987). Threats to internal validity have been identified by Campbell and Stanley (1963) and Bernstein (1976) and include history, maturation, testing, instrumentation, statistical regression, selection bias, and selection-maturation interactions. Although outcome research also should address threats to the external validity of the research, internal validity has a more important impact on research; it is discussed more fully in the following section and by Snow and Tebes (this volume).

The classic approach to outcome research is the randomized control study, which compares and contrasts the intervention effects on individuals randomly assigned to a treatment condition, an attention placebo control, a comparative program, or a treatment-as-usual condition. In many clinical trials, this outcome research model can be enhanced by the use of double-blind procedures in which the subject and program provider are "blind" or are not knowledgeable concerning treatment levels and subject's assignment. Although drug prevention research studies have used randomized controlled experiments with the individual, school, class, or community as the sampling unit and unit of assignment, few if any drug prevention studies have initiated double-blind studies.

As an alternative, quasi-experimental designs have been effectively used when randomization is not feasible. In those instances, Campbell and Stanley have proposed that a variety of quasi-experiments can be considered, including time series designs, nonequivalent control groups, and a variety of separate sample pre-post designs. However, each of these designs has individual strengths and weaknesses related to internal and external validity threats. Flay (1986) suggests that outcome research, which he calls "efficacy" trials, also could use a "historical control" design when randomized control studies are not possible and the comparison groups consist of one or more conditions from a previous trial that are not randomly assigned.

It is important to note that process and outcome research focus on specific program effects and should be included within the same study. Whereas outcome research provides the quantitative measure of program effects, process research provides description of the program's theory, content, and social and cultural context. In addition, operational characteristics are documented, drawing from quantitative and qualitative data to help explain why some program elements may have worked and why other elements may have failed to achieve the predicted change levels. Other chapters in this volume (Snow and Tebes, Bentler, Dwyer and Mackinnon, Biglan and colleagues) elucidate salient issues for drug abuse prevention outcome research.

Impact Research—Assessing the Effectiveness of Preventive Interventions

Impact research is distinctly different from process and outcome research because it may not be program specific. Rather, impact research assesses the cumulative and/or aggregate effects of prevention programs operating within a geographic area and over a specified period (French and Kaufman 1984). The geographic area could be defined as a school or school system, county, town, city, state, region, or the Nation. Impact research attempts to examine the

effectiveness of preventive interventions implemented under real-world conditions.

The purpose of impact research is to measure significant changes in drug-related indicators at the macro or community level and to link them with prevention effects established through process and outcome research of specific prevention programs. Major indicators for impact research include the assessment of trends in the rates of incidence and prevalence of drug use and abuse as measured by community or national epidemiological drug surveys such as the National Household Survey or the National High School Senior Survey on Drug Abuse; drug-related morbidity, including Drug Abuse Warning Network data from hospital emergency rooms; drug-related mortality from medical examiner reports; and drug-related accident data.

A prototypic example of prevention impact research was recently published by Bachman and colleagues (1988). Analyses of data from the National High School Senior Survey on Drug Abuse have repeatedly shown a downward trend in self-reported marijuana use since 1978. For example, the reported use of marijuana within the past 30 days by high school seniors decreased from 37.1 percent in 1978 to 18.0 percent in 1988 (Johnston et al. 1989).

Using a series of multivariate and bivariate analyses relating drug use trends from 1976 to 1986, Bachman and coworkers (1988) assessed several explanations for the downward trend in reported regular marijuana use. This research indicates that although lifestyle factors such as truancy, religious values, or political beliefs are linked to individual differences in the use of marijuana, these factors alone do not explain the general downward trend in marijuana use since 1978. Rather, they report that the 10-year decline in marijuana use by high school seniors was directly related to the increase in the perception of risk of psychological and physical harm as well as the increased perception of personal disapproval associated with the regular use of marijuana reported by high school seniors from 1978 to 1986. For example, the number of high school seniors reporting great risk of personal harm due to smoking marijuana regularly rose from 34.9 percent in 1978 to 71.3 percent in 1986, whereas the number of high school seniors who reported their disapproval of smoking marijuana regularly increased from 67.5 percent in 1978 to 86.6 percent in 1986. Figure 6 illustrates the relationship between the reported use of marijuana within the past 30 days and reported perception of harmful consequences of regular marijuana use (Johnston et al. 1989). Bachman and coworkers (1988) conclude that if the perceived risks of harm and personal disapproval associated with regular marijuana use had not risen substantially since 1978, the decline in use of this drug would not have occurred. The authors suggest that over time, adolescents gradually began to realize that the

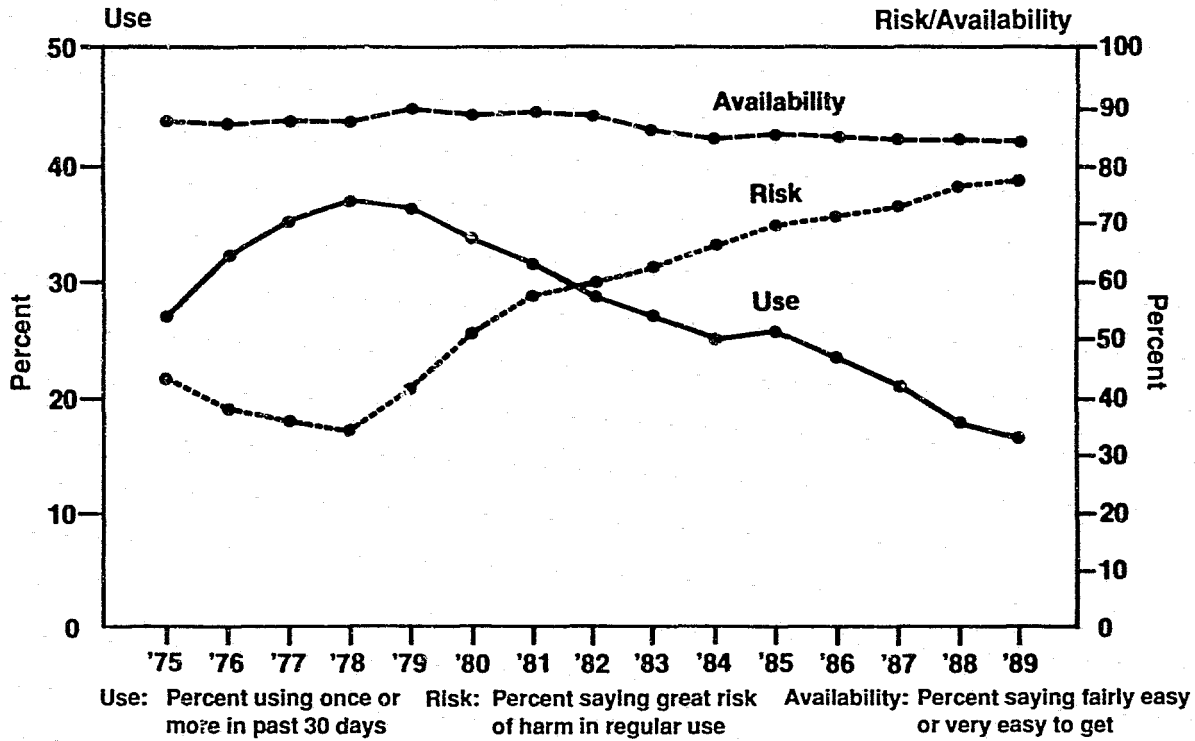

FIGURE 6. *Marijuana 30-day use and perceived harm from the National High School Senior Survey on Drug Abuse, 1975-88*

SOURCE: Johnston et al., "Monitoring the Future" study.

use of marijuana could result in physical and social consequences that were detrimental to one's health and social acceptance. Preventive messages conveyed by the "system" through drug education courses and media broadcasts and particularly by peers concerning the negative effects of marijuana appeared to be consistent with the social experience of high school seniors and may have prompted a significant change in their use of marijuana.

Likewise, the recent downturn in reported cocaine use has been associated with an increased perception of risk of harm resulting from regular use, occasional use, and use of cocaine once or twice. For example, the reported use of cocaine within the past 30 days by high school seniors decreased from 6.7 percent in 1985 to 3.4 percent in 1988, whereas the perception of harm resulting from the use of cocaine once or twice rose from 34 percent in 1985 to 54.2 percent in 1988 (Johnston et al. 1989). As with marijuana, the increase in perception of harmful consequences in reported cocaine use by high school seniors may help explain the recent decline.

Although scientific knowledge of the efficacy of individual prevention programs is still evolving through controlled outcome research, the general downward trends in marijuana use (and more recently cocaine use) suggested by this impact research may have resulted from increased perceived risk of harm and perceived disapproval—two major objectives of drug prevention programs. These findings provide encouraging initial indications that appropriate and scientifically sound drug prevention education and media messages can effectively reach a large proportion of the adolescent population and change drug-related attitudes and behaviors.

The primary methodological concern for the design of impact research is external validity, which refers to the capacity of the experimental research design to generalize to other populations, settings, and times (Campbell and Stanley 1963; Snow, this volume). Factors that threaten external validity or the representativeness of the research include reactive effects of testing that may increase or decrease the participant's response to the intervention; interaction effects between subject selection and the intervention; reactive effects due to the experimental environment that do not carry over to implementing the intervention in nonexperimental settings; effects of multiple treatments administered to the same respondents; reactivity between social and cultural norms and the intervention; and nonsystematic variability in program adoption/adaptation at the local level. Chapters in this volume by Hawkins and colleagues, Forman and Linney, and Schinke and Orlandi discuss important methodological procedures to enhance the external validity of drug abuse prevention research studies.

SUMMARY

Drug abuse preventive intervention research requires a comprehensive metamethodology that yields scientifically sound data to assess the theory, implementation process, efficacy, and effectiveness of drug prevention programs. This chapter recommends the use of a research methodology that systematically focuses on process, outcome, and impact research techniques and procedures. In addition, an example of drug impact research is suggested using National High School Senior Survey on Drug Abuse data.

REFERENCES

Arnold, C.; Kuller, L.; and Greenlick, M., eds. *Advances in Disease Prevention.* Vol. 1. New York: Springer Publishing Company, 1981.

Auslander, G. Social networks and the functional health status of the poor. *J Community Health* 13(4):197-209, 1988.

Bachman, J.; Johnston, L.; O'Malley, P.; and Humphrey, R. Explaining the recent decline in marijuana use: Differentiating the effects of perceived risks, disapproval, and general lifestyle factors. *J Health Soc Behav* 29:92-112, 1988.

Bernstein, I., ed. *Validity Issues in Evaluative Research.* Beverly Hills: Sage Publications, 1976.

Brook, J.; Gordon, A.; Whiteman, M.; and Cohen, P. Dynamics of childhood and adolescent personality traits and adolescent drug use. *Devel Psychol* 22(3):403-414, 1986.

Brown, S.; Stetson, B.; and Beatty, P. Cognitive and behavioral features of adolescent coping in high-risk drinking situations. *Addict Behav* 14:43-52, 1989.

Bry, B. Predicting drug abuse: Review and reformulation. *Int J Addict* 18(2):223-233, 1983.

Bukoski, W. School-based substance abuse prevention: A review of program research. In: Griswold-Ezekoye, S.; Kumpfer, K.; and Bukoski, W., eds. *Childhood and Chemical Abuse: Prevention and Intervention.* New York: Haworth Press, 1986. pp. 93-116.

Bukoski, W. The Federal approach to primary drug abuse prevention and education. In: Inciardi, J., ed. *Handbook of Drug Control in the United States.* Westport, CT: Greenwood Press, 1990.

Bukoski, W. A definition of drug abuse prevention research. In: Donohew, L.; Sypher, H.; and Bukoski, W., eds. *Persuasive Communication and Drug Abuse Prevention.* Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Campbell, D., and Stanley, J. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally College Publishing Company, 1963.

Cloninger, C. Etiologic factors in substance abuse: An adoption study perspective. In: Pickens, R.W., and Svikis, D.S., eds. *Biological Vulnerability to Drug Abuse.* National Institute on Drug Abuse Research Monograph 89. DHHS Pub. No. (ADM)88-1590. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1988. pp. 52-72.

Cook, R. The alternatives approach revisited: A biopsychological model and guidelines for application. *Inter J Addict* 20(9):1399-1419, 1985.

Cook, R.; Lawrence, H.; Morse, C.; and Roehl, J. An evaluation of the alternatives approach to drug abuse prevention. *Int J Addict* 19(7):767-787, 1984.

Dolan, L.; Kellam, S.; Brown, C.; and Laudolff, J. Short term impact of a mastery learning preventive intervention on early risk behaviors. *Am J Community Psychol*, in press.

Flay, B. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 15:451-474, 1986.

French, J., and Kaufman, N. *Handbook for Prevention Evaluation.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)84-1145. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1984.

General Accounting Office. *Drug Abuse Prevention: A Report to the Select Committee on Narcotics Abuse and Control, House of Representatives.* GAO Pub. No. HRD-88-26. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1988.

Glynn, T.J.; Leukefeld, C.G.; and Ludford, J.P., eds. *Preventing Adolescent Drug Abuse: Intervention Strategies.* National Institute on Drug Abuse Research Monograph 47. DHHS Pub. No. (ADM)83-1280. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983.

Goodstadt, M. School-based drug education in North America: What is wrong? What can be done? *J Sch Health* 56(7):278-281, 1986.

Gorsuch, R., and Butler, M. Initial drug abuse: A review of predisposing social psychological factors. *Psychol Bull* 83(1):120-137, 1976.

Hawkins, D.; Doueck, H.; and Lishner, D. Changing teaching practices in mainstream classrooms to improve bonding and behavior of low achievers. *Am Educ Res J* 25(1):31-50, 1988.

Hawkins, J.; Lishner, D.; Catalano, R.; and Howard, M. Childhood predictors of adolescent substance abuse: Toward an empirically grounded theory. In: Griswold-Ezekoye, S.; Kumpfer, K.; and Bukoski, W., eds. *Childhood and Chemical Abuse: Prevention and Intervention.* New York: Haworth Press, 1986. pp. 11-48.

Hawkins, J., and Nederhood, B. *Handbook for Evaluating Drug and Alcohol Prevention Programs.* Alcohol, Drug Abuse, and Mental Health Administration. DHHS Pub. No. (ADM)87-1512. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1987.

Jason, L.; Thompson, D.; and Rose, T. Methodological issues in prevention. In: Edelstein, B., and Michelson, L., eds. *Handbook of Prevention.* New York: Plenum Press, 1986. pp. 1-19.

Johnston, L. The etiology and prevention of substance use: What can we learn from recent historical changes. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 155-177.

Johnston, L.; O'Malley, P.; and Bachman, J. *Drug Use, Drinking, and Smoking: National Survey Results from High School, College, and Young Adult Populations, 1975-1988.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)89-1638. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1989.

Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. "Monitoring the Future Study: A Continuing Study of the Lifestyles and Values of Youth." National Institute on Drug Abuse. University of Michigan, Institute for Social Research, Ann Arbor, MI.

Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Kumpfer, K. Special populations: Etiology and prevention of vulnerability to chemical dependency in children of substance abusers. In: Brown, B., and Mills, A., eds. *Youth at High Risk For Substance Abuse.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)87-1537. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1987. pp. 1-72.

Leukefeld, C., and Moskowitz, J. Discussions and recommendations. In: Glynn, T.J.; Leukefeld, C.G.; and Ludford, J.P., eds. *Preventing Adolescent Drug Abuse: Intervention Strategies.* National Institute on Drug Abuse Research Monograph 47. DHHS Pub. No. (ADM)83-1280. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983. pp. 250-255.

Linsky, A., and Straus, A. *Social Stress in the United States: Links to Regional Patterns in Crime and Illness.* Dover, MA: Auburn House Publishing Company, 1986.

Manatt, M. *Parents, Peer, and Pot II.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)83-1290. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983.

Morehouse, E. Working in the schools with children of alcoholic parents. *Health Soc Work* 4(4):145-162, 1979.

Myrick, R., and Erney, T. *Youth Helping Youth: A Handbook for Training Peer Facilitators.* Minneapolis: Educational Media Corporation, 1979.

National Institute on Alcohol Abuse and Alcoholism. *Alcohol and Health.*

DHHS Pub. No. (ADM)87-1519. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1987.

National Institute on Drug Abuse. *Adolescent Peer Pressure: Theory, Correlates, and Program Implications for Drug Abuse Prevention.* DHHS Pub. No. (ADM)86-1152. Rockville, MD: Supt. of Docs., U.S. Govt. Print. Off., 1981.

National Institute on Drug Abuse. Prevention research. In: *Drug Abuse and Drug Abuse Research: The Second Triennial Report to Congress.* DHHS Pub. No. (ADM)87-1486. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1987. pp. 33-58.

National Institute on Drug Abuse. *Drug Abuse Curriculum for Employee Assistance Program Professionals.* DHHS Pub. No. (ADM)89-1587. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1989.

New Jersey Department of Public Safety. *Attorney General's Statewide Action Plan for Narcotics Enforcement.* Trenton, NJ: New Jersey Department of Law and Public Safety, 1988.

Newcomb, M. *Drug Use in the Workplace: Risk Factors for Disruptive Substance Use Among Young Adults.* Dover, MA: Auburn House Publishing Company, 1988.

Newcomb, M.; Maddahian, E.; and Bentler, P. Risk factors for drug use among adolescents: Concurrent and longitudinal analyses. *Am J Public Health* 76(5):525-531, 1986.

Oetting, E., and Beauvais, F. Peer cluster theory, socialization characteristics, and adolescent drug use: A path analysis. *J Counseling Psychol* 34(2):205-213, 1987.

Office of National Drug Control Policy. *National Drug Control Strategy.* Doc. No. S/N 040-000-00543-9. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990.

Pentz, M.A. Local government and community organizational strategy for drug abuse prevention: Theory and methods. In: Coombs, R., and Ziedonis, D. *Handbook on Drug Abuse Prevention.* Englewood Cliffs, NJ: Prentice-Hall, in press.

Pentz, M.; Dwyer, J.; MacKinnon, D.; Flay, B.; Hansen, W.; Wang, E.; and Johnson, C. A multicommunity trial for primary prevention of adolescent drug abuse. *JAMA* 261(22):3259-3266, 1989.

Pickens, R.W., and Svikis, D.S., eds. *Biological Vulnerability to Drug Abuse.* National Institute on Drug Abuse Research Monograph 89. DHHS Pub. No. (ADM)88-1590. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1988.

Rush, B.R.; Gliksman, L.; and Brook, R. Alcohol availability, alcohol consumption and alcohol-related damage: The distribution of consumption model. *J Stud Alcohol* 47(1):1-10, 1986.

Schaps, E.; DiBartolo, R.; Moskowitz, J.; Palley, C.; and Churgin, S. Primary prevention evaluation research: A review of 127 impact studies. *J Drug Issues* 11:17-43, 1981.

Schaps, E.; Moskowitz, J.; Malvin, J.; and Schaeffer, G. *The NAPA Drug Abuse Prevention Project: Research Findings.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)84-1339. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1984.

Schuckit, M. Biological vulnerability to alcoholism. *J Consult Clin Psychol* 55(3):301-309, 1987.

Schuerman, J. *Research and Evaluation in the Human Services.* New York: The Free Press, 1983.

Shaver, J. The verification of independent variables in teaching methods research. *Educ Researcher* 12(7):3-9, 1983.

Shoemaker, P., ed. *Communication Campaigns About Drugs: Government, Media, and the Public.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1989.

Simons, R.; Conger, R.; and Whitbeck, L. A multistage social learning model of the influences of family and peers upon adolescent substance abuse. *J Drug Issues* 18:293-315, 1988.

Stamler, J. Lifestyles, major risk factors, proof, and public policy. *Circulation* 58:3-19, 1978.

Stamler, J. Public health aspects of optimal serum lipid-lipoprotein levels. *Prev Med* 8:733-759, 1979.

Tobler, N.J. Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *J Drug Issues* 16(4):537-567, 1986.

AUTHOR

William J. Bukoski, Ph.D.
Program Director
Intervention Research
Prevention Research Branch
Division of Epidemiology and Prevention Research
National Institute on Drug Abuse
Rockwall II, Suite 615
5600 Fishers Lane
Rockville, MD 20857

140138

Prevention Intervention Research: Challenges and Opportunities

Richard R. Clayton and Anne Cattarello

INTRODUCTION

We cannot change the past, only interpret and reinterpret it. There are many things that have been done or not done, sins of omission and commission, that all of us wish we had been smart enough to avoid. We therefore need to follow parental advice—confront the past, learn from mistakes, and do not repeat them in the future.

Prevention intervention research is a "new" field. In the United States, the National Institute on Drug Abuse (NIDA) and the National Institute on Alcohol Abuse and Alcoholism have been in existence about 15 years.

At the beginning of any new endeavor, researchers stake out the territory, identify the parameters of the problem they seek to understand, and tackle the most salient research questions. They often are hampered by scarce resources and a small knowledge base; interventions are expensive and labor intensive. This exploratory phase of research is more often hypothesis generating than hypothesis confirming. The methodological errors are sometimes glaring but provide a basis for learning what not to do.

The Prevention Branch at NIDA was not created until 1982 (Bell and Battjes 1985), a bureaucratic acknowledgment that special attention needed to be given to an area of drug abuse research that was beginning to achieve scientific credibility and significance. The progress is impressive. The current relatively strong knowledge base in the prevention intervention field is growing quickly and holds much promise and challenge. However, as in most fields of intellectual and applied endeavor, research has opened new windows of opportunity and created a need to explore new frontiers in the quest to understand the phenomenon of drug use and abuse and to prevent it and its consequences.

KNOWLEDGE DEVELOPMENT IN THE PREVENTION INTERVENTION FIELD: A DESCRIPTIVE MODEL

One of the great shortcomings in the graduate training of sociologists is that no clinical experience is required. Clinical psychologists and psychiatrists receive clinical training but are often shortchanged by not receiving much of a macro-oriented perspective. Also, in the drug abuse field, many "epidemiologists" have no formal training in epidemiology. In a sense, this is a function of the amount of time available, the limited capacity to learn, and the difficulty of teaching a sincere appreciation and respect for other disciplines to impatient "true believers."

Reasons for Choosing Nicotine Distribution as a Model

Nicotine and how it is distributed in the body of a smoker was chosen as a relevant model for describing the development of the knowledge base in prevention intervention research for several reasons.

First, most prevention intervention research is school based and is designed to deter completely or at least delay the onset of drug use. Thus, much of the focus of such research is on the so-called gateway drugs: nicotine, alcohol, and marijuana.

A second reason for using nicotine as a model is that drug abuse prevention research has been strongly influenced by efforts at cardiovascular risk reduction. Smoking is one of the key risk factors for cardiovascular disease. A significant shortcoming of prevention intervention research to date may be its failure to focus directly on drugs. (Note that "smoking," not nicotine, was mentioned.) The emphasis on smoking, a route of administration instead of a drug, may have had significant effects on findings concerning drug abuse prevention.

Third, a biologically based example was chosen because the drug abuse prevention intervention field, as social scientists know it, is as parochial in its orientation to drug abuse as are other fields in which scientists believe that isolating the genetic and other biologically and chemically based markers of vulnerability will "prevent" drug abuse.

Distribution of Nicotine in a Heavy Smoker

When a heavy smoker wakes up in the morning, the blood level of nicotine is low, usually about 5 to 8 ng/mL, well below the comfort zone of about 18 ng/mL. A nanogram is one-billionth of a gram, so these are small amounts of the drug.

The comfort zone is labeled the nicostat and functions much like a thermostat in a home or office. The nicostat is governed by nicotinic receptor sites that constantly monitor the blood for the concentration of nicotine. Therefore, when the heavy smoker wakes up, he or she is nicotine deficient and begins to experience withdrawal symptoms if the drug is not provided.

Generally, the first thing a heavy smoker does in the morning is to smoke several cigarettes, usually closely together, inhaling more deeply than later in the day. When the body is deficient of nicotine, the nicotine from the first cigarettes of the day go into the tissues (Benowitz and Jacob 1984; Henningfield and Jasinski 1988; Russell 1988). Until the tissues have been saturated, the blood level of nicotine cannot increase. The "distributional half-life" of nicotine, when the tissues are not saturated, is 9 minutes. After four half-lives, clinically significant amounts of the drug are no longer present. So, in 36 minutes, virtually all the nicotine from the first cigarette of the day has been distributed and metabolized. This accounts for the number of quick cigarettes early in the morning. After the tissues have become saturated, it is possible to begin seriously raising the blood nicotine level. At this point, the "metabolic half-life" of nicotine is operative. The metabolic half-life of nicotine is 120 minutes; four metabolic half-lives is 8 hours, which explains why most smokers can sleep throughout the night without having to smoke. The nicotine built up during the day is metabolized during the night. First, that which is in the blood is metabolized. Then, when that is gone, the nicotine stored in the tissues begins to flow back into the blood to be metabolized. This explains why the smoker is nicotine deficient early in the morning.

The knowledge base in the prevention intervention field is similar to the example cited above. For some time now, we have been in the initial distributional phase of research: The half-life of findings has been relatively short and somewhat idiosyncratic. This has been an exploratory phase in which the samples have been small, the methodological flaws apparent, and the components mixed and matched to see which things seem to work best.

Although the number of studies conducted has been reasonably large and the articles published impressive, the conclusion reached by Flay is instructive:

> Overall, the findings from the most rigorous studies to date suggest that the social influences approach to smoking prevention can be effective some of the time. However, this conclusion seems to be somewhat fragile, given the considerable differences between studies in the patterns of reported results. Also, at least two plausible alternative interpretations of the reported results remain—namely,

effects of testing (or screening), and the Hawthorne effect . . . the social influences approach to smoking prevention is an efficacious approach, further research is needed on the conditions under which the social influences programs are effective, for whom they are effective, and why they work (Flay 1985, pp. 90-91).

We may still be in the initial distributional phase of knowledge development about prevention intervention, and given the progress to date, the metabolic half-life of prevention intervention research may be within sight (Botvin 1986; Tobler 1986; Schaps et al. 1981).

ASSUMPTIONS ABOUT PREVENTION INTERVENTION RESEARCH: LIMITATIONS FOR THIS PRESENTATION

Anyone seriously interested in drug abuse prevention and cognizant of the history of efforts to evaluate it would endorse the following assumptions:

- Drug abuse prevention interventions are relevant across the entire life cycle.
- Drug abuse prevention interventions must target not only individuals but also families, work groups, the worksite, neighborhoods, entire communities, other organizational contexts, and society at large.
- Drug abuse prevention interventions can be extremely short (a 15-second media message) and unconnected to other prevention messages or can be long and integrated into a coherent, consistent menu of messages or curriculum materials.
- There may be important differences among individuals in their susceptibility to the influence of prevention interventions.
- Given the degree of social concern about drug abuse and the incredible magnitude of efforts focused on prevention interventions, it may be difficult or impossible to sort out the confounds that interfere with valid assessments of program efficacy.
- Because of the inherent logic of primary prevention efforts and the fact that youth are a "captive" audience organized around age-graded activities, most of the exploratory research has been directed at preventing initiation and delaying onset of drug use and has been school based.

With these assumptions in mind, the principal focus of this chapter is on school-based interventions. However, to the extent that research design and measurement issues are generic, the review also is relevant to prevention interventions occurring in other contexts. In addition, special attention is given to preliminary analyses from a community-wide evaluation of Project DARE (Drug Abuse Resistance Education), a primary prevention program offered to sixth graders in 17 lessons and taught by a trained police officer.

OUTCOME, PROCESS, AND IMPACT: AN IMPORTANT TRILOGY

Outcome

Most prevention studies appropriately focus the most attention on outcomes, providing answers to the following types of questions:

- Do the students who received the prevention intervention exhibit "better" rates of use at posttest or followup than those who did not receive the intervention?
- Are there fewer new users in the treatment as opposed to the no-treatment (control) condition?
- Have attitudes become more antidrug in the experimental condition than in the control condition?
- Are students in the treatment schools more knowledgeable about drugs and their effects than their counterparts in the control schools?
- Are the students in the treatment schools more skilled than students in control schools in resisting peer pressure to use drugs?
- Do students in treatment schools more accurately perceive the behavioral norms concerning drug use among their same-age peers than students in the control schools?

These are outcome questions about knowledge, attitudes, perceived norms, peer resistance skills, and drug use and represent the "bottom line" in primary prevention programing: If the changes are not in the desired direction and/or if the differences between those who received the prevention and those who did not are not statistically significant, then the tendency is to claim that the intervention "failed."

Process

However, outcomes are only part of the prevention process. An attribution of "failure" presumes the intervention occurred under controlled conditions. This means at least two things related to process. First, it assumes that all students in the treatment condition received approximately equal exposure to the intervention and that those responsible for applying the treatment were approximately equal in their ability to deliver the intervention. Second, and perhaps most important, it assumes that there is initial equivalence on important variables between those in the treatment and control conditions.

Unfortunately, in most prevention intervention studies there is insufficient control exercised over the prevention process to claim that lack of statistically significant differences means the intervention failed.

Impact

If relatively few studies have exerted enough control over the process of prevention, even fewer have assessed impact. In school-based prevention programs, impact can occur in ways that are extremely important but may not be reflected in an outcome evaluation. For example, the extensive discussion of drug abuse prevention curriculums and programs by school personnel and the commitment to provide a specific program are vital parts of increasing public awareness of and commitment to solving the problem of drug abuse. It is a sufficiently salient topic that may get more parents involved with their community schools.

Prevention programs can have a definite impact on the school milieu. For example, in 1986-87 the DARE program was implemented on a pilot basis in five treatment schools, which were compared to five control schools. In the first full year of implementation, there were 23 elementary schools randomly assigned to receive the DARE program and 8 that served as controls. One of the principals of a school that had been part of the pilot test and whose school was randomly selected into the treatment condition during the first year insisted that the DARE officer be in his school on Mondays. During the pilot phase he had been tracking attendance/absence rates. He said that attendance on Mondays during the pilot year had been significantly higher, particularly among those youth he considered to be at high risk for drug abuse. Further, he reported significantly fewer disciplinary problems on the days the DARE officer was in the school to deliver the prevention curriculum.

DARE officers report that the younger siblings of their previous students come into the class having absorbed much of the information delivered to their older

siblings. The impact of a school prevention program on siblings, parents, and families may be important in addition to whether there are significant differences in knowledge, attitudes, and practices observed between treatment and control conditions.

When seeds are planted, some require time to take root; others take longer. One of the impacts of an elementary school-based program may be delayed changes in the norms at the junior and senior high school levels. If the focus is primarily on outcomes, important effects may be overlooked at longer term followups. By failing to assess impact in addition to process and outcome, the "baby may be thrown out with the bathwater."

Therefore, there are at least three integral elements to assess in school-based prevention programs—outcome, process, and impact. As scientists are inexorably drawn to outcome, so as teachers we should be keenly aware of and interested in process; however, as citizens our ultimate interest may be impact.

TREATMENT VS. CONTROL: BETWEEN-GROUP DIFFERENCES MAY MASK CRUCIAL WITHIN-GROUP DIFFERENCES

Most school-based prevention studies involve random assignment of schools to either treatment or control conditions. When the unit of assignment is the school, the appropriate unit of analysis is the school (Cook and Campbell 1979). However, in most existing studies, using the school as the unit of analysis would be impossible because of the small number of schools involved. Therefore, the most common approach is to use the individual as the unit of analysis and to compare those students in experimental schools with those in control schools on salient outcome variables.

Question 1. Were the Experimental and Control Groups Constituted by Random Assignment?

The model used in most evaluations of prevention interventions is that of the classic experiment. The purpose of random assignment is to neutralize the effects of variables that might produce a spurious interpretation of results. However, randomization is generally easier to implement within a laboratory setting, where most or all of the relevant variables can be controlled, than it is in the field, where many variables cannot be controlled. An implicit assumption of random assignment is that within-group differences are minimized so that the only key difference between experimental and control groups is the intervention.

Sometimes, what occurs is "mostly" random assignment. For example, in one of the most comprehensive prevention intervention studies to date, the Waterloo

Study (Best et al. 1988), matched pairs of schools (on socioeconomic status, size, rural/urban location) were created in two school districts (six pairs in one district and five in another). Flay and colleagues (1985) indicate that "Assignment to treatment or control conditions from the matched groups was random except in three cases . . . where an administrator thought that principals might not cooperate as fully if their schools were assigned to the control condition."

In the Midwestern Prevention Project (Project STAR—Students Taught Awareness and Resistance), the initial universe was 50 schools. However, only 42 made it into the final eligible pool because some had closed or consolidated with other schools (Pentz et al. 1989). Pentz and coworkers report, "Of the remaining 42 schools, 8 were assigned randomly to program or control conditions, 20 could reschedule existing programming and were assigned to the program condition, and 14 did not have the flexibility to reschedule existing programming and were assigned to the control condition." Thus, in one of the largest and most comprehensive (i.e., complex) efforts at community-based prevention, random assignment of schools to receive the STAR curriculum (modeled after the Project SMART curriculum [Hansen et al. 1988]) occurred for only 8 of the 42 schools included in the evaluation.

Random assignment is essential to credible results concerning outcome. Without random assignment, the results of a study may be interesting and provocative but not completely persuasive because of the possible confounds that may render the interpretations spurious (Graham et al. 1984).

Question 2. What Was Randomly Assigned, and How Many Units Were Located in the Treatment and Control Conditions?

The convention is to randomly assign whole schools to either the treatment or control condition. This has the advantage of limiting the confounding that can occur when students in one school are assigned to different conditions. However, it also has a disadvantage. As Flay and colleagues (1985) state:

> When schools are assigned randomly to conditions, and when the program is delivered to intact classes, the school is the most appropriate unit of statistical analysis for some purposes and the classroom is appropriate for others. It is not entirely appropriate to use the individual as the unit of statistical analysis, except for those questions that concern (a) the effects of different levels of attention to or participation in a curriculum on program effectiveness or (b) the effect of the program on individuals with a differential risk to become smokers (e.g., those with smoking vs. nonsmoking parents or friends).

When schools are used as the unit of analysis, most studies run into problems with the size of n. Even when there are large numbers of individuals providing the data, the number of schools may be relatively small. It is here that an intersection of methodological strategies becomes problematic. The experimental design is best implemented where control is possible; field experimental studies of drug abuse prevention interventions often more closely resemble general social surveys.

In a sense, with peer resistance curriculums, the classroom may be the most appropriate unit of analysis because it is in the classroom that resistance skills are practiced. It is at the classroom level that differential exposure to the intervention is most salient. Furthermore, when the intervention occurs at the elementary school level, the levels of use are extremely low. Flay and coworkers (1985) illustrate this by saying, "Having only zero to three children from any one grade level start smoking in a treatment group while five or six do so in a control group cannot lead to any confidence that the finding is robust." In epidemiological studies this is known as a problem of floor effects or asymmetry. The possible confidence interval around a prevalence rate of 5 is considerably less than the possible confidence interval around a rate of 25. In addition, an increase from 5 to 10 is a 100-percent increase while a 5-unit increase when the base is 25 is only a 20-percent increase.

This problem is especially important in some studies where the focus is on the number of "new" users of a substance; those who already have used the substance will be eliminated from analysis. Because drug initiation is, to some extent, a maturational or developmental phenomenon, exclusion of past users significantly reduces the number of persons "at risk" for initiation and may provide unrealistic comparisons.

Question 3. Was There Initial Equivalence Between the Experimental and Control Groups on Relevant Variables?

One test of the effectiveness of random assignment is to compare the experimental and control groups on important variables at pretest. If the groups are initially equivalent, barring the intrusion of significant events over which the researcher can exert no control, the only "real" difference between the groups should be the intervention. Testing for initial equivalence is absolutely essential. This should include not only sociodemographic variables such as race/ethnicity, sex, age, presence of siblings at home, and family structure but also baseline prevalence rates for use of the substances the curriculum or intervention is designed to prevent. Helping young people never to start using drugs is only part of the goal; delaying onset of use is another; getting those who have started not to progress to more regular use or to other drugs is still

another. In terms of overall impact on the health of a community and the Nation, getting those who are already regular users of drugs to reduce or stop consumption may be the most important target of prevention (i.e., those most at risk for abuse of drugs are those who start using gateway drugs the earliest).

Initial equivalence is important for the presumed etiologic variables as well. If one posits a strong predictive relationship between self-esteem and drug use and there is clear nonequivalence between experimental and control groups on this predictor, attributions of the efficacy of the intervention tied to the influence of self-esteem would be inappropriate. The same is true of attitudes.

Initial equivalence on presumed predictor variables may be a mixed blessing. Many prevention interventions experience a ceiling effect. Elementary school-age children generally see things in black-and-white terms; gray is not a color of choice. They hold views with intensity, so gradations of intensity are often not present. Therefore, experimental and control groups may be initially equivalent on attitudes toward smoking, for example. However, the mean scores on this variable for both groups may be so high and the standard deviation so constrained that the attitudinal scale is essentially a constant and any changes from baseline to posttest may be nothing more than random noise.

Any reports of prevention interventions should contain a full-blown evaluation of initial equivalence, indications of how data were transformed if the groups were not initially equivalent on some variables, and full disclosure on problems with floor and ceiling effects. The ultimate goal of science is replicability. The only way to simultaneously reduce type I and type II errors is to replicate the study.

Question 4. Were There Important Differences in How the Intervention Was Delivered?

Teachers are different from each other. Classes are different and force the teachers to behave somewhat differently, even when the content of the material is the same. Teachers often have other responsibilities besides teaching a drug abuse curriculum. Therefore, they may shave time from the drug abuse curriculum to spend more time teaching other "more important" subjects.

Any attempt to attribute differences between treatment and control groups to the intervention must be able to assert that the intervention was delivered faithfully, exactly the way it should be delivered. One of the purposes of process evaluation is to examine fidelity to the curriculum. Another is to examine the influence of the context of the teaching situation on the delivery of the curriculum (Perry et al. 1983; Perry et al. 1986; Botvin et al. 1983).

In the DARE curriculum, the teaching is conducted by police officers who have received 80 hours of intensive instruction concerning teaching strategies and the specific lessons included in DARE. Police officers in the United States are taught to follow protocols rigorously and to conduct their behavior "by the book." In addition, these police officers do not assign grades. If there is little variation between the officers delivering the DARE curriculum, then we have eliminated one of the sources of confounds for interpreting the outcomes and impacts of DARE.

PRELIMINARY DATA ON DARE IN LEXINGTON: CHECKING THE POSSIBLE CONFOUNDS

Pilot Study

In the 1986-87 school year, a decision was made by the Lexington-Fayette County Kentucky Police Department and the school system to implement a primary drug abuse prevention program in the sixth grade (almost all sixth graders are 11 years old). The prevention program and curriculum were developed by the Los Angeles Police Department in 1984 and consist of 17 lessons lasting about 1 hour each. It is taught by police officers in uniform but without a gun who undergo 80 hours of training to deliver this program. The police department and the school system agreed to conduct a pilot study involving the five schools that had already been selected as experimental schools and to allow researchers to collect data from five control schools matched to the experimental schools.

In spite of matching instead of randomization, initial equivalence was achieved on all of the salient sociodemographic variables. There were a few significant differences between experimental and control schools, with differences in the expected direction (i.e., favorable toward DARE) with one exception: The students in the control schools that received the drug abuse prevention component of the health curriculum had stronger antismoking attitudes than did the students in the DARE pilot schools. This was a difference that was anticipated because of the relative attention given to cigarette smoking in DARE in comparison with the existing health curriculum.

In addition, the response of the teachers, principals, and students to the program was uniformly enthusiastic. The mayor and police chief, who were initially skeptical and distant, noticed the positive press and enthusiasm generated by the program and began to change their attitude toward police officers being taken off the beat and put into the classroom for "demand reduction" instead of "supply reduction" activities. This is an important potential impact of this particular prevention program.

Politics and Science

A policy decision was made to fully implement the DARE curriculum throughout the county. However, to facilitate understanding of the long-term effects of this program, the police department and the school system agreed to the random assignment of 23 schools to DARE for the 1987-88 school year with 8 control schools waiting until the 1988-89 school year to receive the DARE program. The study reported here is thus a modified cohort sequential design with experimental and control schools present only in the first year, with the school as the unit of random assignment.

With regard to prevention interventions, most decisions to implement a particular program usually are made without benefit of scientific research or data. A problem is perceived; a decision is made to do something about it; and resources are committed to solving the problem. The police department and school system had "political" decisions to make. They were sophisticated enough to recognize the need for research on the efficacy of the program but sensitive enough to the issue of drug use among children to make decisions in lieu of compelling positive or negative evidence about the efficacy of DARE.

Prevention intervention research almost always involves tempering scientific needs for rigor with political and other realities. The real world is not neatly partitioned into experimental and control groups.

Initial Equivalence

The purpose of testing for initial equivalence (table 1) is twofold: (1) to verify that the process of randomization evened out any relevant differences between treatment and control groups and (2) to identify any relevant places where preintervention differences might make spurious interpretations about changes in experimental and control groups.

In the first cohort of sixth graders exposed to the DARE curriculum in 23 schools and the sixth graders who received the standard drug abuse component in the 8 control schools, the number of items on which significant differences occurred was relatively small.

Demographically, there was a significant difference between treatment and control schools with regard to race: The treatment schools had a larger percentage of white students while the control schools had a relatively larger percentage of black students. Although the schools are neighborhood based, busing is used to achieve as much racial balance as possible. The data from the Monitoring the Future studies of high school seniors (Johnston, in press) are

TABLE 1. *Initial equivalence for 23 treatment and 8 control schools: Evaluation of Project DARE in Lexington, Kentucky*

Items and Scales for Assessing Initial Equivalence	Treatment (N=1,434)	Control (N=485)
Gender		
Male	51.3%	50.3%
Female	48.7%	49.7%
Race (chi-square=8.66, df=2, p<.01)		
White	77.3%	71.0%
Black	19.2%	25.5%
Other	3.5%	3.5%
Birth cohort		
1974 or before	8.4%	9.1%
1975	39.2%	40.2%
1976	51.5%	50.0%
1977 and after	.8%	.6%
Number of siblings		
None	17.3%	17.9%
One or more	82.7%	82.1%
Attitudes items and scales*		
How well doing in school	1.84	1.84
How often attend church	2.80	2.85
How important is church	3.01	3.10
How happy most of time	1.77	1.77
Do not need drugs to feel good	1.53	1.60
Anyone who uses drugs belongs in jail	2.36	2.34
Okay to buy alcohol if get away with it	4.74	4.82 (.01)**
Cigarettes not harmful if few are used	4.23	4.32
Kids who smoke can quit any time	3.86	4.00
Kids who drink are more grown up	4.68	4.77 (.01)
Okay kids drink if quit before habit	4.32	4.47 (.01)
Safe to take another's prescription	3.91	3.92
Cops rather catch you than help	4.18	4.04
Okay to use many aspirin with headache	4.54	4.62
Okay ride with drinker if seems not drunk	4.65	4.67
If parents use drugs, must be okay	4.80	4.83
PCP causes strange behavior	2.03	2.02
Other scales		
Scale of self-esteem	39.36	39.41
Scale of peer relationships	30.34	30.93
Scale of family relationships	11.49	12.48 (.001)
Scale of general attitudes toward drugs	32.41	32.88
Self-reports of drug use		
Lifetime use of cigarettes	28.3%	29.0%
Past year use of cigarettes	15.0%	13.0%
Past month use of cigarettes	4.7%	5.0%
Lifetime use of alcohol	32.0%	26.0% (00.6)
Past year use of alcohol	20.4%	15.3% (00.1)
Past month use of alcohol	10.6%	5.4% (.001)
Lifetime use of marijuana	4.1%	6.1%
Past year use of marijuana	3.1%	1.6%
Past month use of marijuana	3.1%	1.3%
Lifetime use of smokeless tobacco	14.7%	15.0%
Past year use of smokeless tobacco	7.8%	7.5%
Past month use of smokeless tobacco	4.0%	2.7%

*1=strongly agree, 2=agree, 3=neutral, 4=disagree, 5=strongly disagree
**Numbers in parentheses indicate statistically significant differences between treatment and control groups.

consistent: There was underreporting of drug use by racial and ethnic minorities. Therefore, any initial lack of equivalence should serve to buffer rates of reported drug use in the control schools.

The attitude items ranged from scores of 1 (strongly agree) to 5 (strongly disagree). Only three of the individual attitude items produced a statistically significant difference between treatment and control groups, all of which dealt with alcohol use and all of which showed students in the control group to be less positive toward alcohol. These differences may be reflected in the higher prevalence rates for alcohol use in the treatment vs. the control schools. Finally, there were statistically significant differences with regard to the scale measuring family relationships (control group less positive toward family) and the scale measuring general attitudes about drugs (control schools more negative toward drugs).

It is safe to conclude that on most of the salient variables, there was initial equivalence between treatment and control groups. Care must be exercised in interpreting results that include those items, scales, and prevalence rates where initial equivalence was not present and where race might have operated as a confound to definitive interpretations.

It should be noted that this test for initial equivalence is at a macrolevel (treatment vs. control groups) and not at the level at which random assignment occurred (the school level). Therefore, although initial equivalence can be asserted, there may be rather large differences between schools "within" the treatment and "within" control conditions that could confound interpretations of efficacy.

Process Evaluation

The DARE officers were asked to designate their "best," "regular," and "worst" classes. At least two process evaluators attended these classes to ensure that the officer delivered the same lesson (i.e., content) in each class.

Officers were rated with regard to the following elements: (1) mastery of the material, (2) compliance with the lesson plan, (3) coverage of all aspects of the lesson, (4) clarity of the communication of the material, (5) clearness of speaking, (6) use of audiovisuals, (7) participation by students, (8) extent to which students were serious about the class, (9) behavior of students during the lesson, (10) evidence of rapport between the officer and students, (11) extent of interaction between the officer and students, and (12) the free exchange of ideas within the classroom. Ratings were from 1 to 5.

There are several important findings. First, on each item for each class, each officer received ratings above the middle (i.e., higher than 3). In fact, the majority of ratings were near 5. Simply put, the evaluators felt that the officers were good teachers. Second, there was almost no difference among the officers on any of the elements on which they were evaluated. Finally, the only difference worthy of note occurred for "behavior exhibited by the students." In each instance, although the evaluators were blind to the officer's perception of the class (i.e., best, regular, worst), the worst classes were much lower on behavior than the other classes. This suggests that classroom context may be an important variable to consider in prevention intervention studies.

Outcome Evaluations

In most school-based interventions, the hypothesis is that there will be fewer new users in the treatment than in the control condition. The data in table 2 show that although there were differences in new users of cigarettes and marijuana across all 31 schools in this study (i.e., baseline before intervention to posttest 16 to 19 weeks later), there were no statistically significant differences between experimental and control schools in the percentage of new users of these substances (cigarettes, smokeless tobacco, alcohol, marijuana).

TABLE 2. *Initiation of use of drugs from baseline to posttest: Evaluation of Project DARE in Lexington, Kentucky*

		Initiation of Drug Use, Baseline to Posttest			
Schools and Officers	N	Smokeless Cigarettes	Tobacco	Alcohol	Marijuana
All Schools	31	14.2%*	7.6%	11.5%	3.2%*
Treatment Schools	23	14.7%	8.1%	11.5%	3.5%
Control Schools	8	12.7%	6.1%	11.3%	2.2%
Officer 1	4	16.4%	7.5%	2.2%	9.0%
Officer 2	8	14.7%	8.0%	14.5%	4.4%
Officer 3	8	13.7%*	8.3%	13.7%	2.8%**
Officer 4	3	15.1%	8.7%	11.8%	5.9%

*Statistically significant differences:
All schools, cigarettes, chi-square=44.7, df=30, p<.04
All schools, marijuana, chi-square=55.1, df=30, p<.002

**Statistically significant differences across schools for:
Officer 3, cigarettes, chi-square=25.6, df=7, p<.001
Officer 3, marijuana, chi-square=16.3, df=7, p<.02

As noted earlier, the process evaluation revealed fidelity to the curriculum and virtually no differences among the police officers delivering the DARE curriculum. The data in table 2 show that there were no statistically significant differences across the four officers with regard to new users of the four substances. However, for Officer 3, there were significant differences across the schools in which he taught in new users of cigarettes and new users of marijuana.

Because the school was the unit of random assignment in this study, school-level data should be examined. For cigarettes, the data in table 3 show that failure to note "significant" differences between treatment and control schools could be a problem of the baseline prevalence rate and the available "pool of eligibles." The baseline prevalence rate for cigarettes ranged from 10 percent in one school for Officer 3 to 57 percent in a school for Officer 2.

TABLE 3. *Use of cigarettes among sixth-grade students in Project DARE in Lexington, Kentucky: Baseline rate, number eligible for initiation, new users, and posttest lifetime prevalence rate*

	Lifetime Experience With Cigarettes				
Officers and Schools	Initial N	Baseline Rate	Number Eligible for Initiation	New Users	Posttest Lifetime Prevalence Rate
Officer 1	282	27%	207	34 (16%)	109 (39%)
School 1	68	24%	52	17%	37%
School 2	34	47%	18	28%	62%
School 3	129	27%	103	13%	30%
School 4	51	33%	34	20%	47%
Officer 2	404	31%	278	41 (15%)	167 (41%)
School 5	34	26%	25	28%	47%
School 6	59	31%	41	15%	41%
School 7	61	21%	48	17%	34%
School 8	19	26%	14	7%	34%
School 9	63	57%	27	18%	65%
School 10	71	17%	59	7%	22%
School 11	33	48%	17	24%	61%
School 12	64	27%	47	13%	36%
Officer 3	514	24%	388	53 (14%)	179 (35%)
School 13	29	10%	26	27%	34%
School 14	94	11%	84	10%	19%
School 15	24	25%	18	6%	29%
School 16	88	18%	72	26%	40%
School 17	56	43%	32	19%	54%
School 18	94	35%	61	3%	37%
School 19	30	17%	25	20%	33%
School 20	99	29%	70	7%	34%
Officer 4	174	32%	119	18 (15%)	73 (42%)
School 21	61	34%	40	10%	41%
School 22	68	21%	54	20%	37%
School 23	45	44%	25	12%	51%

Thus, in school 13, 90 percent of the students were eligible for initiation of cigarette use, whereas in school 9 only 43 percent of the students were eligible for onset of use of cigarettes. Consequently, if the emphasis of the curriculum is on preventing the onset or delaying the onset of use of cigarettes, for almost 6 of 10 students in the latter school the message may be irrelevant.

Similar data for onset of use of marijuana can be seen in table 4. In 5 of the 23 treatment schools, none of the sixth-grade students reported having tried marijuana. The highest baseline rate of marijuana use occurred for Officer 3, 14.5 percent. By posttest, only 1 of the 23 treatment schools had no students reporting marijuana use and 1 school in which 22.8 percent of the sixth graders had tried marijuana by the end of the sixth grade. One implication of these data is that primary prevention perhaps should start in earlier grades.

TABLE 4. *Use of marijuana among sixth-grade students in Project DARE in Lexington, Kentucky: Baseline rate, number of eligible for initiation, new users, and posttest lifetime prevalence rate*

		Lifetime Experience With Marijuana			
Officers and Schools	Initial N	Baseline Rate	Number Eligible for Initiation	New Users	Posttest Lifetime Prevalence Rate
Officer 1	286	3.4%	276	6 (2.2%)	16 (5.6%)
School 1	68	2.9%	66	0.0	2.9%
School 2	36	11.1%	32	6.2%	16.7%
School 3	130	0.8%	129	0.8%	1.5%
School 4	52	5.7%	49	6.1%	11.5%
Officer 2	407	5.4%	385	17 (4.4%)	39 (9.6%)
School 5	35	11.4%	31	12.9%	22.8%
School 6	60	5.0%	57	7.1%	11.7%
School 7	63	3.2%	61	3.3%	6.3%
School 8	18	5.6%	17	0.0	5.6%
School 9	63	11.2%	56	7.2%	17.5%
School 10	69	0.0	69	0.0	0.0
School 11	34	11.7%	30	3.3%	14.7%
School 12	65	1.5%	64	3.2%	4.6%
Officer 3	520	2.9%	505	14 (2.8%)	29 (5.6%)
School 13	28	0.0	28	3.6%	3.6%
School 14	91	0.0	91	3.3%	3.3%
School 15	28	0.0	28	14.3%	14.3%
School 16	89	1.1%	88	2.2%	3.4%
School 17	55	14.5%	47	2.1%	16.4%
School 18	100	2.0%	98	1.0%	3.0%
School 19	30	3.3%	29	3.4%	6.7%
School 20	99	3.0%	96	1.0%	4.0%
Officer 4	177	4.0%	170	10 (5.9%)	17 (9.6%)
School 21	59	5.1%	56	9.0%	13.6%
School 22	71	0.0	71	1.4%	1.4%
School 23	47	8.5%	43	9.3%	17.0%

Another implication is methodological in nature. As in the classic book *Animal Farm*, all animals are created equal, but some are more equal than others. Random assignment of a relatively small number of units into two groups, in this case schools, does not guarantee that homogeneity exists (i.e., within-group variance completely taken care of), only that it should be minimized.

Although schools are the unit of assignment in this and most other studies, the key element in peer-resistance training occurs in the classroom where students practice those skills. Moreover, children are not randomly assigned to classrooms. Regardless of what the administrative regulations may state as ideal, some principals review the roster of fifth graders and decide to keep the "problem kids" together to concentrate the trouble and social control efforts. Another principal may decide to spread the problem kids out across the sixth-grade classrooms to defuse the peer influence process.

Although clearly a simplified version of what probably occurs, the data in table 5 show the lifetime prevalence rates of cigarette use in the classrooms taught by Officer 3 at posttest. In the second school, the rates of cigarette use at baseline are quite similar across the classrooms. In the fifth school, 67 percent of the students in one classroom have tried cigarettes by the end of the sixth grade compared with approximately 30 percent in each of the other two classrooms.

TABLE 5. *Use of cigarettes by the end of sixth-grade among students taught by Officer 3 by classroom by school: Project DARE in Lexington, Kentucky*

	Have Tried Cigarettes by Posttest			
Schools for Officer 3	Classrooms			
	1	2	3	4
School 13	21.4%	37.5%		
School 14	23.1%	16.7%	19.0%	16.0%
School 15	33.3%	18.7%		
School 16	42.9%	37.5%	40.0%	47.6%
School 17	67.3%	28.0%	31.2%	
School 18	39.1%	26.9%	39.5%	41.7%
School 19	28.6%	31.2%		
School 20	25.9%	34.5%	50.0%	29.2%

These data illustrate the point that classroom, school, and neighborhood context may be important variables to measure in addition to data obtained from individual students. These macrolevel or "aggregate"-level variables may be as or even more robust in accounting for behavior and behavior change than the typical psychosocial risk factors employed in most prevention intervention evaluations (Hawkins and Catalano 1989; Clayton 1989; Brunswick 1989; Brook et al. 1988).

PREVENTION INTERVENTION RESEARCH: SCHOOL-BASED PROGRAMS; OTHER, MORE SPECIFIC CONSIDERATIONS

The article by Flay and coworkers (1985) is one of several reviews of the existing literature (Botvin 1986; Best et al. 1988; Schinke and Gilchrist 1985). These reviews provide an excellent overview of the major methodological issues and flaws in the area of prevention intervention research. A descriptive approach, rather than a detailed review, appears below.

Implementation Issues

1. *Different curriculums* are used by different investigators. Although curriculums are probably similar across investigators to date, no systematic review compares the degree of overlap to uniqueness. Thus, claims of efficacy assume that those programs classified as social skills/ social influences are sufficiently similar to be so classified.

2. *Different pedagogical strategies* are used to communicate these curriculum materials, and different media are used in the presentation.

3. *Skills at different levels of abstraction or specificity* are targeted in these curriculums because the programs are based on different assumptions about the appropriate level of abstraction/skills required. Botvin and Eng (1980, 1982) and Botvin and colleagues (1980) focus primarily on "generic" skills, whereas the attitude-behavior and behavioral intentions literature would suggest targeting behavior-specific skills.

4. *Youth at different developmental levels* are the recipients of these curriculums. There are marked and dramatic changes that occur in children (cognitively, socially, psychologically) between the fifth and the eighth grades (Dielman et al. 1986). Furthermore, the way the school system is organized (middle school vs. a more traditional junior and senior high model) may have an influence on the context within which these curriculums are delivered.

5. *Exposure to the prevention intervention* seems to differ significantly across the studies (i.e., number of sessions and spacing between sessions). For example, Project DARE involves 17 lessons whereas the Midwestern Prevention Project (Project STAR) curriculum, which is modeled after and similar to the curriculum for Project SMART (Hansen et al. 1988), involves only 10 sessions.

6. *Presenters of the curriculum occupy different roles* (i.e., teacher, researcher, other adult, peer) *vis-a-vis* the recipients, which could have an impact on the way the program is perceived and received.

7. *The ability of presenters to communicate the program* may differ in important ways that have been ignored. This may occur even when the presenters receive the same training. It may have less to do with formal skills taught or learned than with innate talents and abilities.

8. *The degree to which the curriculum is embedded in the school milieu and infused into the school day* may be an important factor in success. If students do not receive a formal grade for participation in the prevention program and are taught by someone other than "the" teacher, they may be more receptive and attentive.

9. *Measure of exposure may not be connected to measured outcomes.* Although the number of sessions is obviously important, process-oriented measures of participation and amount learned may be a more important measure.

10. *Drug prevention occurring in other institutional contexts* within which these youth must operate—family, church, community organization—may be a factor in success. Not much is being done to assess the influence of the family, church, other organizations, peer groups outside of home, the neighborhood, or community or the extent to which school peers overlap with nonschool peers in influencing onset or delay of onset of drug use.

Design, Measurement, and Analysis Issues

1. *Size of n* in each condition can be a problem. The usual approach is to emphasize the number of students while ignoring or failing to discuss the number of units assigned. For example, in the Midwestern Prevention Project, Pentz and colleagues (1989) gathered data from more than 5,000 students. However, the size of n is 42 when school is the unit of analysis.

2. *Random assignment* most often is done at the school level. This is acceptable but creates a problem if all schools in a community are not included in the sampling frame or if there is a small number of schools. Then, in most cases, the schools assigned to the treatment condition(s) are assumed to be homogeneous, as are the control schools. In several studies, there has been an attempt first to match pairs of schools on presumed salient characteristics (race/ethnicity, socioeconomic status, etc.) and then to randomize into treatment and control conditions.

   This general approach is sound in theory, but then all classes are grouped into the condition to which the school was assigned. There is often as much variability within schools as there is across schools. In addition, teachers and students are not randomly assigned. Some classrooms have populations that may receive the treatment but are considerably more prepared cognitively to be affected by it than are students in another class.

   For example, students are assigned to the sixth grade by the principal. One principal may look over the roster of fifth graders and make a decision to put all of the troublemakers into one classroom for the sixth grade. Another principal may decide to spread the trouble and hope that the "infection" rate will be dissipated. Thus, a management decision for purposes of social control may help explain large differences between "within" classrooms on important variables.

3. *No treatment control groups may be receiving prevention.* It usually is assumed that the control group is essentially not receiving prevention when it is the specific curriculum being evaluated that they are not receiving. In the pilot evaluation of Project DARE in Lexington, the no-treatment control group (i.e., non-DARE) had stronger antismoking attitudes than the DARE students, probably because the section on smoking was stronger in the regular science/health curriculum than it was in DARE.

4. *The number of treatment conditions can complicate the analysis* strategies by varying so many presumed factors (peer vs. teacher, number of sessions, etc.) that it is virtually impossible to make definitive statements about what was delivered and what the effects were (McCaul and Glasgow 1985).

5. *Length of followup* is crucial if the goal is to deter drug use or delay onset because it is important that all persons have passed through the principal "ages of vulnerability" before this is assessed (problem of censored variable).

6. *Attrition* effects are important because at some point an assumption has to be made that those who are not measured at subsequent points are like those who remained in the study and provided data at each point (Biglan et al. 1987; Hansen et al. 1985).

7. *Data collectors and their involvement with program delivery* may be important. If the person gathering the data, usually via a self-administered questionnaire, is also the person who delivers the program, there is the possibility of expectancy effects. If someone else is gathering the data, there is the possibility of lack of trust and rapport.

8. *Data collection for important measured variables.* Paper-and-pencil tests alone, as opposed to supplementary material from observation of class and class members plus, for example, biological testing, are important considerations. It may be important to ask if paper-and-pencil test items are appropriately worded for the age group answering the items and whether the items have been used previously with persons this age. Also, what are the psychometric properties of the scales used in the study? Issues of readability, fatigue, lack of understanding, different reading levels within a classroom, and the cultural sensitivity of the items need to be addressed.

9. *Classroom contextual factors* that are physical and social in nature may influence the results: How many students in what space at what time of the day and under what climate conditions? Sociometrically, what are the factors that may influence answers?

10. *Components delivered and whether they are "balanced"* or weighted in terms of amount of time devoted to them are important considerations. Prevention programs are usually multifocused and involve attempts at disseminating information, influencing attitude formation and change, teaching decisionmaking skills, and enhancing self-esteem. The balance among these different components as well as the perceived connectedness among them may be crucial.

11. *Those already smoking at baseline may be the most important from the perspective of drug abuse.* Those already using the drug are ignored in many studies or given only slight attention in other studies. Using cigarette smoking as an example, the pretest smokers can change status from baseline to posttest to followup status and from smoker to nonsmoker. For example, at three points in time, McNeill and coworkers (1989) studied nicotine intake in girls 11 to 14 years of age by use of saliva cotinine concentrations. The data show some important transitions

that occur in smoking status for adolescent girls. Nonsmoking or even smoking status at an early age is not always predictive of later status. Using point prevalence rates as a measure leads to some inconsistencies in whether one would be labeled a "success" or a "failure."

Of perhaps even more importance from a drug abuse perspective is the degree to which these young women were receiving significant amounts of nicotine. McNeill and colleagues (1989) report that:

> . . . these girls were receiving substantial doses of nicotine from a very early stage in their smoking careers. This suggests that pharmacological effects of nicotine were already important in perpetuating their smoking. Other analyses of the 1985 data indicate an early development of pharmacological addiction; young smokers reported experiencing calming effects when smoking and withdrawal effects during attempts to give up which were similar to those experienced by adult smokers.

This suggests the possibility of using paper-and-pencil tests of nicotine dependence such as the Horn-Russell or the Fagerstrom scales for those who report smoking. Chemical verification may be more important as a measure of the pharmacological impact on the participants than as a way of improving validity of self-reports (Evans et al. 1977).

12. *Clarification is needed on what we are trying to prevent.* The overwhelming focus on smoking probably reflects the influence of cardiovascular risk reduction research, not only on research but also on the researchers involved with drug abuse prevention. It also may reflect the fact that most drug abuse prevention researchers come from developmental/cognitive psychology and not drug abuse research backgrounds. The failure to include other delivery systems for nicotine is a major shortcoming. Fortunately, some researchers are beginning to include smokeless tobacco as a major dependent/outcome variable.

DRUG ABUSE PREVENTION INTERVENTION: AIMING FOR THE BIGGER PICTURE

Using nicotine and its distribution as a model of knowledge development in the prevention intervention field, the focus begins on prevention or delay of onset of

use of the gateway drug nicotine (Clayton et al. 1988; Clayton 1989; Henningfield et al. 1990), particularly use of cigarettes (Evans et al. 1977). There are several reasons for this focus: (1) Nicotine is often the first drug with which young people experiment; (2) it is easily detectable in expired air carbon monoxide samples or in saliva thiocyanate, thus providing a validity check on self-reports of use; and (3) cigarette smoking is identified as one of the key risk factors for cardiovascular disease. The drug abuse prevention intervention field is built on that field of research.

In only 15 years, significant progress has been made in methodologically evaluating the efficacy of prevention interventions for cigarette smoking and, to a lesser extent, the use of alcohol, marijuana, and cocaine. Most of the progress has been made in evaluating school-based programs that are designed around development of cognition and skills that will allow students to resist pressures (perceived and real) to use these drugs. Even with such progress, many of the methodological confounds have not been examined systematically. Although this field of research may be close to entering the "metabolic half-life" phase of research, it is not there yet.

However, the next phase already has been entered. Interventions with multiple components targeted at the entire community (Pentz et al. 1989) is that next phase. The perceived magnitude of the drug problem is resulting in more expansive efforts in spite of the fact that this leads to an almost geometric expansion of the confounds for a clear interpretation of the results and potential efficacy of the interventions.

In evaluating attempts to prevent drug use and abuse, particularly at the community level and in settings where the majority of recipients of the intervention are from mainstream America, there is an important temptation to avoid. Existing epidemiological data from general population surveys consistently show secular trends downward for use of most drugs (Johnston, in press; Bachman and Johnston 1987). Because researchers look for "statistically significant" reductions in onset and prevalence in the groups receiving the intervention, the tendency is to attribute "good" results to "our" prevention intervention when the results merely may reflect trends and methodological confounds already existing or introduced during the course of the study.

Furthermore, there is an increasing recognition that primary and school-based prevention interventions must be supplemented by other interventions aimed at families and the worksite. Focusing on youth was a good beginning, but a significant proportion of illicit drug abuse in the United States occurs among persons 18 to 34 years of age (DuPont and Clayton 1989). Therefore, new

growth areas for the drug abuse prevention intervention field in the 1990s will be adults, families, the elderly, specific high-risk groups such as delinquents and probationers, those who live in neighborhoods and environments saturated with drugs, and entire communities, using a comprehensive attack on the known risk factors for drug abuse (Hawkins and Catalano 1989; Clayton 1989; Brunswick 1989). Meetings with school and drug treatment personnel suggest an even more challenging potential high-risk group: children born to mothers who have abused drugs during pregnancy. These children start life in a drug-impaired condition, and there is increasing evidence that catching up with their cohort is an unlikely outcome.

Unfortunately, the "science" of prevention interventions expands systematically, inductively, arithmetically (1+1+1+, etc.), while immediate demands for prevention interventions require a geometric expansion (2+2=4+4=8, etc.). The focus of science is on rigorous evaluation and presentation of all the blemishes. Prevention advocates often want unequivocal statements that a prevention intervention "works." Assertions that interventions "work" most often are based on "logic," "intuition," and "hope." It is the role of the scientist evaluator to administer a strong dose of rigorous scrutiny to those assertions.

In the drug abuse prevention intervention field, significant progress has been made in the evaluation of school-based programs, and the methodology is improving. However, the demand for interventions is significantly greater than the supply of interventions that have been subjected to rigorous scrutiny. Although many *opportunities* exist and will continue to exist, the *challenge* will be to change the mandate from:

READY . . . FIRE . . . AIM!
to
READY . . . AIM . . . FIRE!

REFERENCES

Bachman, J.G., and Johnston, L.D. "Drug Use in America: Different Kinds of Change, Different Causes." 1987 University Senior Research Scientist Lecture, University of Michigan, November 9, 1987.

Bell, C.S., and Battjes, R., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Benowitz, N.L., and Jacob, P. Daily intake of nicotine during cigarette smoking. *Clin Pharmacol Ther* 35:499-504, 1984.

Best, J.A.; Thomson, S.J.; Santi, S.M.; Smith, E.A.; and Brown, E.S. Preventing cigarette smoking among school children. *Ann Rev Public Health* 9:161-201, 1988.

Biglan, A.; Severson, H.; Ary, D.; Faller, C.; Gallison, C.; Thompson, R.; Glasgow, R.; and Lichtenstein, E. Do smoking prevention programs really work? Attrition and the internal and external validity of an evaluation of a refusal skills program. *J Behav Med* 10:159-171, 1987.

Botvin, G.J. Substance abuse prevention research: Recent developments and future directions. *J School Health* 56:369-374, 1986.

Botvin, G.J., and Eng, A. A comprehensive school-based smoking prevention program. *J School Health* 50:209-213, 1980.

Botvin, G.J., and Eng, A. The efficacy of a multicomponent approach to the prevention of cigarette smoking. *Prev Med* 11:199-211, 1982.

Botvin, G.J.; Eng, A.; and Williams, C.L. Preventing the onset of cigarette smoking through life skills training. *Prev Med* 9:135-143, 1980.

Botvin, G.J.; Renick, N.L.; and Baker, E. The effects of scheduling format and booster sessions on a broad spectrum psychosocial approach to smoking prevention. *J Behav Med* 6:359-379, 1983.

Brook, J.S.; Whiteman, M.; Nomura, C.; Gordon, A.S.; and Cohen, P. Personality, family and ecological influences on adolescent drug use: A developmental analysis. *Adv Alcohol Subst Abuse* 5:123-161, 1988.

Brunswick, A.F. Young black males and substance use. In: Gibbs, J.T.; Brunswick, A.F.; Connor, M.E.; Dembo, R.; Larson, T.E.; Reed, R.J.; and Solomon, B., eds. *Young, Black, and Male in America: An Endangered Species.* Dover, MA: Auburn House Publishing, 1989. pp. 166-187.

Clayton, R.R. "High Risk Youth and Drug Abuse: A Comprehensive Risk Reduction Strategy." Special report prepared for the Office of Juvenile Justice and Delinquency Prevention, 1989.

Clayton, R.R.; Voss, H.L.; and LoSciuto, L.A. Gateway drugs: What are the stages people go through in becoming drug abusers. *Pharm Times* 53:28-35, 1988.

Cook, T.D., and Campbell, D.T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally, 1979.

Dielman, T.E.; Shope, J.T.; Butchart, A.T.; and Campanelli, P.C. Prevention of adolescent alcohol misuse: An elementary school program. *J Pediatr Psychol* 11:259-282, 1986.

DuPont, R.L., and Clayton, R.R. "Users of Illicit Drugs: The New Focus of Demand Reduction." Unpublished manuscript, 1989.

Evans, R.I.; Hansen, W.B.; and Mittelmark, M.B. Increasing the validity of self-reports of behavior in a smoking-in-children investigation. *J Appl Psychol* 62:521-523, 1977.

Flay, B.R. Psychosocial approaches to smoking prevention: A review of findings. *Health Psychol* 4:449-488, 1985.

Flay, B.R.; Ryan, K.B.; Best, J.A.; Brown, K.S.; Kersell, M.W.; D'Avernas, J.R.; and Zanna, M.P. Are social psychological smoking prevention programs effective? The Waterloo Study. *J Behav Med* 8(1):37-59, 1985.

Graham, J.W.; Flay, B.R.; Johnson, C.A.; Hansen, W.B.; and Collins, L.M. Group comparability: A multiattribute utility measurement approach to the use of random assignment with small numbers of aggregated units. *Eval Rev* 8:247-260, 1984.

Hansen, W.B.; Collins, L.M.; Malotte, C.K.; Johnson, C.A.; and Fielding, J.E. Attrition in prevention research. *J Behav Med* 8:261-275, 1985.

Hansen, W.B.; Johnson, C.A.; Flay, B.R.; Graham, J.W.; and Sobel, J. Affective and social influences approaches to the prevention of multiple substance abuse among seventh grade students: Results from Project SMART. *Prev Med* 17(2):135-154, 1988.

Hawkins, J.D., and Catalano, R. "Risk and Protective Factors for Alcohol and Other Drug Problems: Implications for Substance Abuse Prevention." Special report prepared for the Office of Juvenile Justice and Delinquency Prevention, 1989.

Henningfield, J.E.; Clayton, R.R.; and Pollin, W. Involvement of tobacco in alcoholism and illicit drug use. *Br J Addict* 85:279-292, 1990.

Henningfield, J.E., and Jasinski, D.R. Pharmacologic basis for nicotine dependence. In: Pomerleau, O.F., and Pomerleau, C.S., eds. *Nicotine Replacement: A Critical Evaluation.* New York: Alan R. Liss, 1988. pp. 35-62.

Johnston, L.D. Toward a theory of epidemics of illicit drug use. In: Donohew, L.; Sypher, H.; and Bukoski, W., eds. *Persuasive Communication and Drug Abuse Prevention.* Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

McCaul, K.D., and Glasgow, R.E. Preventing adolescent smoking: What we have learned about treatment construct validity. *Health Psychol* 4(4):361-387, 1985.

McNeill, A.D.; Jarvis, M.J.; Stapleton, J.A.; West, R.J.; and Bryant, A. Nicotine intake in young smokers: Longitudinal study of saliva cotinine concentrations. *Am J Public Health* 79:172-175, 1989.

Pentz, M.A.; Dwyer, J.H.; MacKinnon, D.P.; Flay, B.R.; Hansen, W.B.; Wang, E.Y.I.; and Johnson, C.A. A multicommunity trial for primary prevention of adolescent drug abuse. *JAMA* 261:3259-3266, 1989.

Perry, C.L.; Klepp, K.; Halper, A.; Hawkins, K.G.; and Murray, D.M. A process evaluation study of peer leaders in health education. *J School Health* 56:62-67, 1986.

Perry, C.L.; Telch, M.J.; Killen, J.; Burke, A.; and Maccoby, N. High school smoking prevention: The relative efficacy of varied treatments and instructors. *Adolescence* 18:561-566, 1983.

Russell, M.A.H. Nicotine replacement: The role of blood nicotine levels, their rate of change, and nicotine tolerance. In: Pomerleau, O.F., and Pomerleau, C.S., eds. *Nicotine Replacement: A Critical Evaluation.* New York: Alan R. Liss, 1988. pp. 63-94.

Schaps, E.; DiBartolo, R.; Moskowitz, J.; Palley, C.; and Churgin, S. Primary prevention evaluation research: A review of 127 impact studies. *J Drug Issues* 11:17-43, 1981.

Schinke, S.P., and Gilchrist, L.D. Preventing substance abuse with children and adolescents. *J Consult Clin Psychol* 53:596-602, 1985.

Tobler, N.J. Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *J Drug Issues* 16(4):537-567, 1986.

ACKNOWLEDGMENT

Work on this study was supported by National Institute on Drug Abuse grant DA-05312.

AUTHORS

Richard R. Clayton, Ph.D.
Professor of Sociology and Psychiatry
Scientific Director

Anne Cattarello, M.A.
Study Director
Project DARE Evaluation

Center for Prevention Research
University of Kentucky
147 Washington Avenue
Lexington, KY 40506-0402

Contributions of Drug Epidemiology to the Field of Drug Abuse Prevention

Lloyd D. Johnston

INTRODUCTION

The interface between epidemiology and drug prevention always has been an extremely important one, and etiologic studies—insofar as they are any different from epidemiological studies—are important to both fields. The dichotomy between epidemiology and etiology is largely an artificial one, deriving from the medicalization of the field of social science research in the drug area and, thus, the arbitrary division into two discrete segments. There are few major epidemiological studies that do not speak to the determinants of use as well as to the quantification of use.

There are at least eight ways in which epidemiological studies should, and often do, inform the development of drug prevention programs and the evaluations of such programs. They can inform prevention efforts by providing a dynamic assessment of the following:

1. Drug use or drug-related problems that need to be prevented

2. Ages at which such use is initiated or problems are occurring

3. Subgroups in the population most "at risk" in terms of their demographic and lifestyle characteristics

4. Changing backdrop against which the effects of specific prevention efforts should be assessed

5. Importance of certain key intervening variables such as attitudes and beliefs

6. Behavioral and moral norms with regard to drug use among young people and other groups having influence on them (information that can be used in designing persuasive messages)

7. Extent to which major classes of prevention programing are reaching targeted segments of the population and the subjective opinions of those populations as to the helpfulness and effects of the interventions

8. Combined effectiveness of all forces in the society that tend to reduce drug use or abuse, including those that are planned programs, more spontaneous efforts of groups or individuals, and other historical events

In all these cases, emphasis has been placed on the dynamic nature of epidemiology's influence because the social realities in each are likely to change. The remainder of this chapter discusses in more detail, under each of these eight areas, the way that epidemiological research has and can influence the formulation of prevention programs and the research designed to evaluate them.

TYPES OF DRUG USE OR DRUG-RELATED PROBLEMS TO BE PREVENTED

Clearly, the mix of illicit drugs used by the American population during the past 20 years has changed dramatically and continually (Johnston et al. 1989; National Institute on Drug Abuse 1988). LSD was one of the major drugs of concern during the early 1970s, as was methamphetamine. As epidemiological data, much of them from our own studies (Johnston et al. 1989), began to document the rapid rise in daily marijuana use among young people, its use became an issue of central concern. The increased concern stimulated social action that contributed to several prevention activities: Increased research attention was given to the drug by Federal scientific agencies; a federally initiated task force was eventually formed at the Institute of Medicine of the National Academy of Sciences to summarize what was known about the effects of marijuana; and the media began to pay more attention to marijuana in their specials and regular news programing. Finally, schools began to use the new knowledge about effects in their curriculums, and more recently, media antidrug advertising campaigns have been using this institution in their antidrug spots.

Epidemiological studies subsequently showed a strong decline in marijuana use in the 1980s (particularly, heavy use), and its relative importance began to fade (table 1 and figure 1). The popularity of PCP rose rapidly in the late 1970s and fell just as fast (figure 2). Cocaine's popularity also rose in the late 1970s, and it remained at peak levels during the first half of the 1980s (figure 3). It is clearly the drug of greatest concern today. As these drugs have risen and fallen in popularity, broadly defined prevention efforts have changed their foci somewhat, and as the epidemiological research and monitoring systems show

TABLE 1. *Trends in annual prevalence of 18 types of drugs by high school seniors*

Percent who used in past 12 months

Approximate N =	Class of 1975 9,400	Class of 1976 15,400	Class of 1977 17,100	Class of 1978 17,800	Class of 1979 15,500	Class of 1980 15,900	Class of 1981 17,500	Class of 1982 17,700	Class of 1983 16,300	Class of 1984 15,900	Class of 1985 16,000	Class of 1986 15,200	Class of 1987 16,300	Class of 1988 16,300	Change 1987-88
Marijuana/hashish	40.0	44.5	47.6	50.2	50.8	48.8	46.1	44.3	42.3	40.0	40.6	38.8	36.3	33.1	-3.2ss
Inhalants[a]	NA	3.0	3.7	4.1	5.4	4.6	4.1	4.5	4.3	5.1	5.7	6.1	6.9	6.5	-0.4
Inhalants adjusted[b]	*NA*	*NA*	*NA*	*NA*	*8.9*	*7.9*	*6.1*	*6.6*	*6.2*	*7.2*	*7.5*	*8.9*	*8.1*	*7.1*	*-1.0*
Amyl and butyl nitrites[c,b]	NA	NA	NA	NA	6.5	5.7	3.7	3.6	3.6	4.0	4.0	4.7	2.6	1.7	-0.9s
Hallucinogens	11.2	9.4	8.8	9.6	9.9	9.3	9.0	8.1	7.3	6.5	6.3	6.0	6.4	5.5	-0.9s
Hallucinogens adjusted[d]	*NA*	*NA*	*NA*	*NA*	*11.8*	*10.4*	*10.1*	*9.0*	*8.3*	*7.3*	*7.6*	*7.6*	*6.7*	*5.8*	*-0.9*
LSD	7.2	6.4	5.5	6.3	6.6	6.5	6.5	6.1	5.4	4.7	4.4	4.5	5.2	4.8	-0.4
PCP[c,b]	NA	NA	NA	NA	7.0	4.4	3.2	2.2	2.6	2.3	2.9	2.4	1.3	1.2	-0.1
Cocaine	5.6	6.0	7.2	9.0	12.0	12.3	12.4	11.5	11.4	11.6	13.1	12.7	10.3	7.9	-2.4sss
"Crack"[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4.1	4.0	3.1	-0.9s
Other cocaine[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	9.8	7.4	-2.4ss
Heroin	1.0	0.8	0.8	0.8	0.5	0.5	0.5	0.6	0.6	0.5	0.6	0.5	0.5	0.5	0.0
Other opiates[e]	5.7	5.7	6.4	6.0	6.2	6.3	5.9	5.3	5.1	5.2	5.9	5.2	5.3	4.6	-0.7s
Stimulants[e]	16.2	15.8	16.3	17.1	18.3	20.8	26.0	26.1	24.6	NA	NA	NA	NA	NA	NA
Stimulants adjusted[e,f]	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*20.3*	*17.9*	*17.7*	*15.8*	*13.4*	*12.2*	*10.9*	*-1.3s*
Sedatives[e]	11.7	10.7	10.8	9.9	9.9	10.3	10.5	9.1	7.9	6.6	5.8	5.2	4.1	3.7	-0.4
Barbiturates[e]	10.7	9.6	9.3	8.1	7.5	6.8	6.6	5.5	5.2	4.9	4.6	4.2	3.6	3.2	-0.4
Methaqualone[e]	5.1	4.7	5.2	4.9	5.9	7.2	7.6	6.8	5.4	3.8	2.8	2.1	1.5	1.3	-0.2
Tranquilizers[e]	10.6	10.3	10.8	9.9	9.6	8.7	8.0	7.0	6.9	6.1	6.1	5.8	5.5	4.8	-0.7
Alcohol	84.8	85.7	87.0	87.7	88.1	87.9	87.0	86.8	87.3	86.0	85.6	84.5	85.7	85.3	-0.4
Cigarettes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

NOTES: Level of significance of difference between the two most recent classes: s=.05, ss=.01, sss=.001.

NA = data not available.

[a]Data based on four questionnaire forms (N=four-fifths of N indicated).
[b]Adjusted for underreporting of amyl and butyl nitrites. See text for details.
[c]Data based on a single questionnaire form (N=one-fifth of N indicated).
[d]Adjusted for underreporting of PCP.
[e]Only drug use that was not under a doctor's orders is included here.
[f]Based on the data from the revised question, which attempts to exclude the inappropriate reporting of nonprescription stimulants.
[g]Data based on a single questionnaire form in 1986 (N=one-fifth of N indicated) and on two questionnaire forms in 1987 (N=two-fifths of N indicated).
[h]Question text changed slightly in 1987.

SOURCE: Johnston et al. 1989.

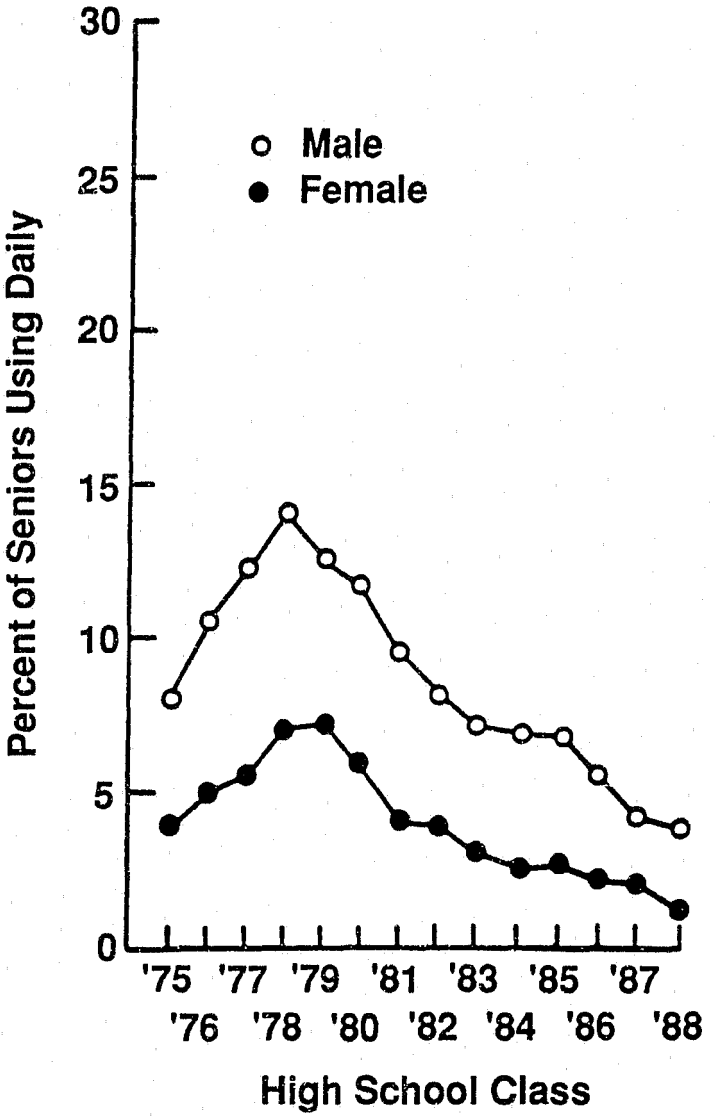

FIGURE 1. *Trends in 30-day prevalence of daily use of marijuana by sex*

SOURCE: Johnston et al. 1989.

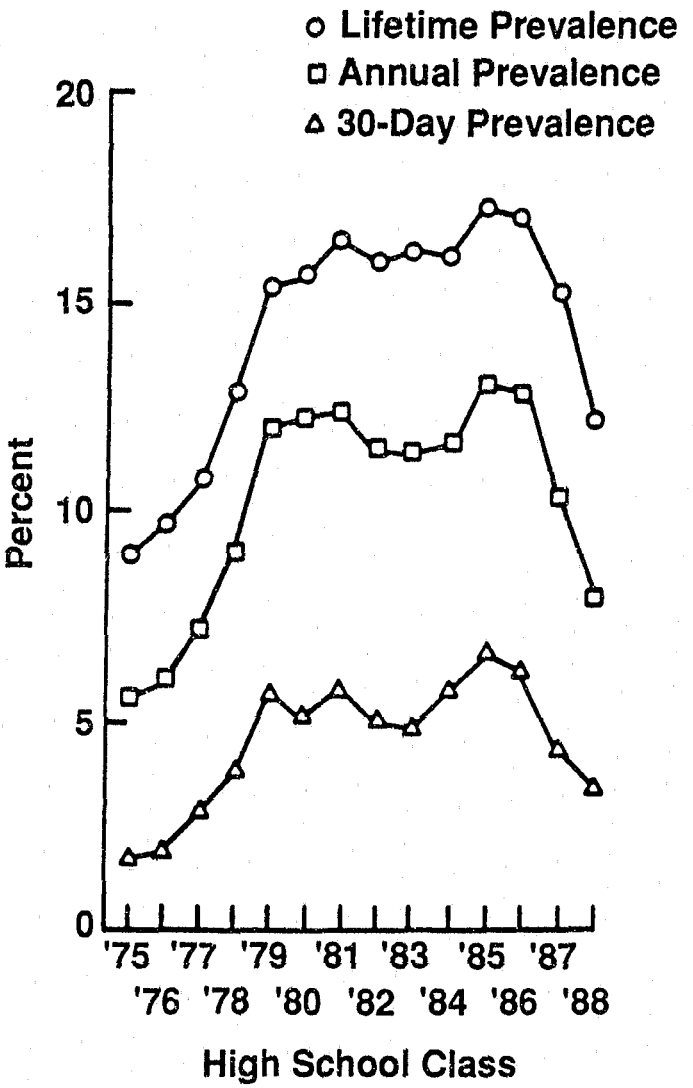

FIGURE 2. *Trends in lifetime, annual, and 30-day prevalence of PCP use—all seniors*

SOURCE: Johnston et al. 1989.

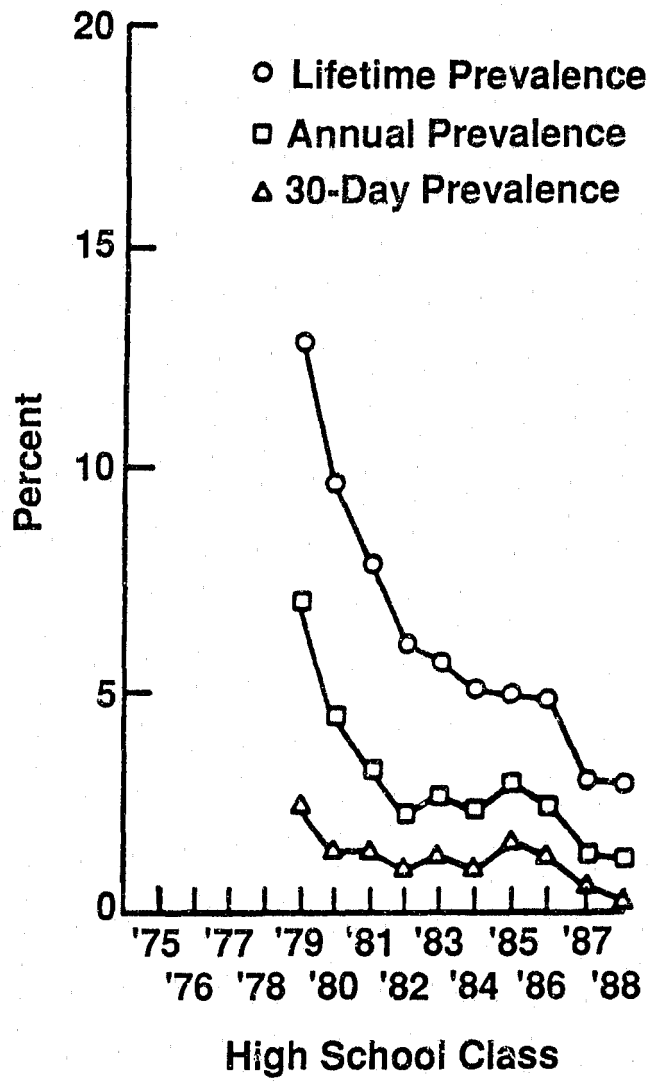

FIGURE 3. *Trends in lifetime, annual, and 30-day prevalence of cocaine use—all seniors*

SOURCE: Johnston et al. 1989.

further developments (Will "ecstasy" be next?), the focus of prevention will continue to change.

A DYNAMIC ASSESSMENT OF THE AGES AT WHICH USE OR PROBLEMS ARE OCCURRING

In the early stages of this epidemic, illicit drug use evolved largely among American college students, before radiating to age peers and downward to high school and, eventually, junior high school students (Johnston 1973; Johnston et al. 1989; Kandel 1978; National Institute on Drug Abuse 1988). Clearly, the ages at which intervention seems appropriate have changed. There is always a concern about stimulating interest and awareness, where perhaps none had existed, by the "premature" introduction of prevention programs; thus, knowing the age progression is important in designing prevention programs. Today, it is clear that many youngsters initiate illicit drug use during junior high school, and they initiate smoking and drinking, with which illicit drug use is highly correlated, earlier still (table 2). Thus, the appropriate ages for intervention are now very young indeed.

The mention of alcohol and cigarettes also points up another area where epidemiology has contributed to prevention design—namely, by identifying the sequential way in which most youngsters proceed into illicit drug use (Johnston 1973; Kandel 1975a). Although the causal influences of the earlier steps on reaching the later ones may not be fully pinned down, there are many in the field who believe there is some causation involved. To the extent that they are right, the importance of preventing smoking and alcohol use at still earlier ages is heightened even further above what it would be just to prevent the considerable adverse consequences that result directly from the use of those drugs.

DEFINING SUBGROUPS IN THE POPULATION MOST AT RISK

Aside from subgroups defined in terms of age, those defined on a host of other dimensions have been identified in epidemiological research as being at higher risk than others, thus helping to target early intervention efforts at, for example, those who drop out of school (Johnston 1973), those in school but with frequent absences or poor grades (Bachman et al. 1981; Johnston 1973; Johnston et al. 1989; Kandel 1975b), those most deviant in other ways (Jessor and Jessor 1977; Johnston 1973; Osgood et al. 1988), those frequently out of the parental home in the evening and those with little attachment to religion (Bachman et al. 1981). Interestingly, although the subgroups most at risk on these behavioral and lifestyle dimensions theoretically could change, they have for the most part remained the same in recent years (Bachman et al. 1986). However, one lifestyle dimension that was once relevant for defining at-risk groups—being a

TABLE 2. *Grade of first use for 16 types of drugs by the senior class of 1988 (percentages)*

Grade in which drug was first used	Marijuana	Inhalants[a]	Amyl/Butyl Nitrites	Hallucinogens[a]	LSD	PCP	Cocaine	Heroin	Other Opiates	Stimulants[a] (adjusted)	Sedatives	Barbiturates	Methaqualone	Tranquilizers	Alcohol	Getting Drunk	Cigarettes	Cigarettes (daily)
6th	2.3	2.4	0.2	0.1	0.1	0.1	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.2	8.6	3.3	19.4	1.5
7th-8th	8.8	3.0	0.3	0.7	0.6	0.3	0.7	0.2	1.1	3.2	1.3	1.2	0.7	1.1	21.9	13.5	19.5	4.2
9th	13.2	3.4	0.7	1.8	1.5	0.7	1.6	0.3	1.8	5.2	2.4	2.1	1.2	2.2	25.7	20.6	11.7	5.3
10th	10.1	2.8	0.6	2.3	2.0	0.8	3.0	0.2	2.6	5.0	1.7	1.4	0.5	1.8	18.2	16.2	7.3	4.2
11th	8.5	2.8	0.8	2.4	2.2	0.6	4.1	0.2	1.8	4.1	1.3	1.2	0.5	2.7	12.0	12.1	5.8	3.5
12th	4.3	2.2	0.6	1.5	1.3	0.3	2.5	0.2	1.1	2.0	0.9	0.7	0.3	1.3	5.6	5.6	2.6	2.1
Never used	52.8	83.3	96.8	91.1	92.3	97.1	87.9	98.9	91.4	80.2	92.2	93.3	96.7	90.6	8.0	28.8	33.6	79.3

NOTE: This question was asked in two of the five forms (N=approximately 6,000), except for inhalants, PCP, and the nitrites, which were asked about in only one form (N=approximately 3,000).

[a]Unadjusted for known underreporting of certain drugs.
[b]Based on the data from the revised question, which attempts to exclude the inappropriate reporting of nonprescription stimulants.

SOURCE: Johnston et al. 1989.

part of the counterculture--is no longer relevant because that dimension has faded in saliency as one on which individuals can even be located (Johnston 1973; Johnston et al. 1987).

There also have been some interesting changes in relative risk for subgroups defined on demographic dimensions. For example, a study by Johnston and colleagues (1989) demonstrated that, over the course of the cocaine epidemic, young people in the Northeast and West developed a considerably higher risk for cocaine use compared with those in the Midwest and South. In the more recent decline phase, the regional rates are beginning to converge again (figure 4). These studies also have shown that the epidemic of crack use, which

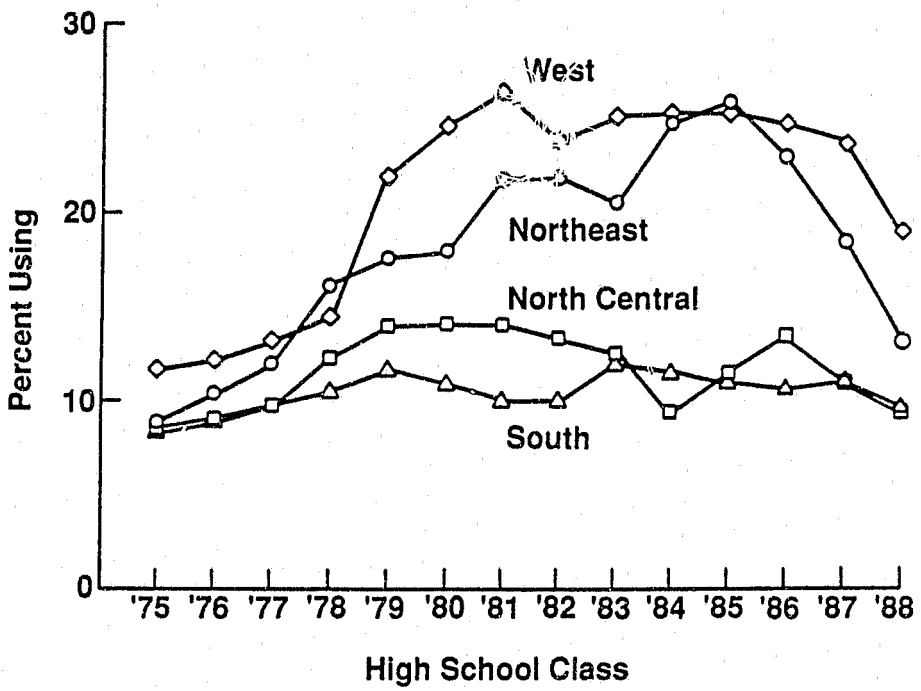

FIGURE 4. *Trends in seniors' lifetime prevalence of cocaine use by region of the country*

SOURCE: Johnston et al. 1989.

started in a few major cities, quickly spread to reach the majority of communities in the country (Johnston et al. 1989).[1] Such findings can and should have a considerable influence on the content and emphasis of prevention programs in different regions and types of communities.

Hopefully, more epidemiological studies focused on inner-city youth (many of whom are dropouts) will begin to evolve, so that we will know with more certainty who is most at risk as well as something about why. It seems that the conditions of living in a poor, inner-city environment are sufficiently different from those faced by American youngsters in the mainstream—about whom we know much more—that the nature and focus of preventive interventions may need to be quite different in those environments. Thus, both quantitative and qualitative epidemiological studies are sorely needed in these communities.

Before leaving the discussion of risk factors for illicit drug use, let us return to the importance of licit drug use as quite visible "markers" of high risk. The powerful association between licit and illicit drug use has been well demonstrated (Miller et al. 1983; Johnston 1987; Kandel 1975a). The one characteristic of youngsters in early adolescence most predictive of risk for eventual illicit drug use probably would be cigarette smoking. The extent to which cigarette smoking has been used in identifying high-risk youth for focused early interventions is unclear, but early use of licit drugs could be used.

ASSESSING THE CHANGING BACKGROUND AGAINST WHICH PREVENTION PROGRAMS MUST BE EVALUATED

Another way in which epidemiological research assists the development and evaluation of prevention is by providing information relevant to answering the question "What would have happened in the absence of the intervention, that is, in the control condition?" A program implemented in the mid-1970s could have had an effect but still have shown no decrease in drug use in the study population because the general population was experiencing an increase in many forms of drug use at that time. In other words, the effects of the program could have been masked or offset by the upward secular trend. Conversely, a program in the early 1980s showing some decline in drug use in a study population may actually have had no effect, because most forms of drug use also were declining in the general population during this period (Johnston et al. 1989).

So far, the issue of age-related change in use has been finessed, but if before-after measures on the same panel of young people are being used to assess the intervention effects, controlling for the normal age-related increment in use is necessary. (With the age ranges that usually are targeted, this probably

would be an increase in use for all drugs, with the possible exception of inhalants.) Clearly, a randomly matched group design is the most rigorous way for assessing what would have happened without the intervention, but such designs are rare and may have very small numbers of groups due to cost constraints. In these cases, ongoing national (and perhaps State) survey series can provide a very rough approximation of what would have been expected in the absence of a particular program. If one is using the national data sets, trends for the relevant region may be a little more appropriate than the overall national trends, but it should be remembered that sampling error is considerably higher for regions. One also might make a multivariate prediction of expected change from the national data sets using several relevant demographic variables to more closely "match" the study population.

One limitation in the national school surveys has been the absence of data on lower grade levels compared with the senior year in high school. This makes specific estimates of expected change for lower grade levels nearly impossible. However, because nearly all changes in drug use (with the primary exception of cigarette smoking) have been shown to be secular trends—that is, trends that are common across age groups (Johnston et al. 1989; O'Malley et al. 1988)—the direction and sharpness of the likely background changes can be estimated. If the national household surveys of drug use (National Institute on Drug Abuse 1988), which track youngsters age 12 and older, are conducted with greater frequency in the future, they may provide more age-relevant change measures for younger age groups. That they use household rather than school interviews likely affects the prevalence estimates obtained, given that greater confidentiality can be obtained in school, but *trend* estimates may not be adversely affected by that methodological difference.

It should be clear that I am not proposing the use of the national survey series as a more desirable, or even close to equivalent, alternative to rigorous experimental designs with comparison groups. I personally believe, and have so testified before Congress, that we need a great many more of the latter. What I am saying is that, in the absence of well-controlled studies, the survey series can be used to deduce very roughly what might have been likely to have happened during that historical period in the absence of intervention.

IDENTIFYING KEY INTERVENING VARIABLES

An additional way in which epidemiological (including etiologic) research can and should facilitate the design and the evaluation of preventive interventions is by helping to identify key intervening variables or factors, such as attitudes, beliefs, and norms, that are important determinants of drug-using behaviors. Insofar as they can be identified, these factors may be targeted in the intervention efforts and probably should be measured in the assessment phase.

It seems that one of the most important theoretical contributions of the "Monitoring the Future" study has been to show that the degree of risk perceived to be associated with a drug can be an important determinant of young people's use of that drug (Bachman et al. 1988; Johnston 1982, 1985; Johnston et al. 1987, 1989). In fact, this has precipitated a first statement of a more general theory of drug use epidemics in which perceived risk plays a central role (Johnston, in press). Our conclusions about the importance of perceived risk were dramatically opposed to the conventional wisdom of the prevention field, which concluded from early findings that "scare tactics" did not work and, therefore, that perceptions of danger were relatively unimportant. It now seems that was a case of "throwing out the baby with the bathwater." The problem with the early efforts to caution youngsters about the dangers of drugs was that they were not credible—that is, they probably did not succeed in changing the perceptions of risk to oneself. Part of the problem may have been the times, as these efforts were made in the early 1970s when young people were not open to cautions from their elders or from "the system"; when many of those cautions were exaggerated or patently false; and, particularly, when many were aimed at forms of drug use that had become symbolic of the youthful counterculture movement. (Recall that the title of the first report of the National Commission on Marijuana and Drug Abuse [1972] was *Marijuana: A Signal of Misunderstanding.)* Perhaps the importance of perceived risk was overlooked in the 1970s, in part, because of the mood of youth in those times and, in part, because the *particular* program implementations aimed at influencing these beliefs were not effective.

This last point leads to a brief aside. It appears that the prevention field may have been too quick to dismiss theoretical notions because particular programs failed to show the desired results. The notion of "alternatives," for example, still seems promising—particularly for youngsters who have a dearth of constructive activities readily available. Yet, it is fair to say that conventional wisdom in the prevention field is that the alternatives approach is a dead issue. It is important that the validity of specific incarnations of an approach not be confused with the validity of the theoretical notion underlying it.

PROVIDING MEASURES OF THE STATISTICAL AND MORAL NORMS AMONG PEERS AND OTHER INFLUENTIAL GROUPS

Because there appears at times to be an exaggerated notion among young people of the proportion of their peers or other significant reference groups taking particular drugs (or accepting the use of such drugs), it has been argued that one way in which data from epidemiological studies can be used in the design of prevention programs is by challenging such instances of "collective ignorance" (Johnston 1985). In other words, data on the actual behavior and

norms of such influential groups can be fed back to the population targeted for prevention.

It is quite clear that the tobacco industry has tried in its advertising to portray smoking as more widely practiced and more widely accepted (particularly in more educated circles) than it is. In other words, the tobacco industry has contributed to an increase in collective ignorance. In addition to working to ban such insidious and misleading advertising, it seems that those in the prevention field might take a page from the book of this exceptionally effective industry and work toward lowering collective ignorance about a host of drugs. Although the case may be clearest for using data collected from youngsters who constitute the target group (e.g., in their own grade or classroom) or a corresponding group of older peers (seniors in their own school), data from broader populations of students and older peers also might be used effectively.

There may be a serious misperception on the part of young people concerning the extent of use and acceptance of drugs in other important groups of role models such as athletes, musicians, and radio and television performers (Johnston 1989). If this concept is true, it could be that a correction of those misperceptions using epidemiological data gathered from those populations would help to change the attitudes, norms, and resolve of young people.

ASSESSING THE COVERAGE AND RATINGS OF WHOLE CLASSES OF PREVENTION PROGRAMING

Quite a different function can be served by epidemiological survey studies when they are used to determine what proportion of an age group (or some other subgroup) is being exposed to certain types of prevention programing and how effective they judge that programing to have been. For example, since 1975, seniors in the Monitoring the Future study have been asked whether they had received in-school drug prevention activities and, if so, of what sort. They also have been asked to rate those experiences. Tables 3 and 4 provide selected findings. Such information is relevant to assessing both the extent of coverage and the aggregate evaluation of the intended audiences. Recalled exposure to school prevention curriculums is probably lower than most people would guess and, although the qualitative ratings show a slight improvement since 1975, there certainly is considerable room for more.

Similar questions in more recent years have been asked about exposure to and judged impact of media campaigns aimed at deterring drug use. Tables 5 and 6 provide some of the results related to those questions, showing a high rate of recalled exposure and, overall, very positive ratings of impact. Roughly two-thirds of seniors recall being exposed to antidrug advertisements weekly or

TABLE 3. *Trends in exposure to drug use prevention elements in school curriculums (percentages)*

	High school seniors in the class of:												
	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
2E15. Have you had any drug education courses or lectures in school?													
1. No	15.7	18.0	20.7	21.0	26.1	23.52	6.22	5.62	7.32	3.92	3.82	1.41	8.9
2. No, and I wish I had	5.1	3.8	4.5	4.7	5.6	4.2	6.0	4.3	3.4	4.3	3.9	4.0	4.1
3. Yes	79.2	78.3	74.8	74.4	68.3	72.4	67.8	70.1	69.2	71.7	72.3	74.6	77.0
N=	2,494	2,556	3,000	2,700	2,710	2,990	2,975	2,719	2,688	2,703	2,568	2,686	2,740
Asked only of those having drug education courses or lectures													
2E17. How many of the following drug education experiences have you had in high school? (Mark all that apply.)													
A. A special course about drugs	22.7	24.8	24.7	22.8	20.5	22.3	20.2	21.4	23.7	20.6	24.1	22.1	23.0
B. Films, lectures, or discussions in one of my regular courses	75.7	74.6	74.7	77.7	76.3	76.8	75.5	77.1	78.0	76.2	77.4	75.1	74.7
C. Films or lectures, outside of my regular courses	28.8	28.2	25.5	22.3	21.0	23.9	25.2	23.9	26.8	30.0	30.4	36.6	40.2
D. Special discussions ("rap" groups) about drugs	24.7	24.1	25.1	22.1	22.4	20.8	20.7	21.2	21.3	19.1	22.5	25.9	22.0
N=	1,979	1,984	2,227	1,980	1,820	2,141	1,987	1,897	1,841	1,929	1,840	1,977	2,095

SOURCE: Johnston et al., "Monitoring the Future" study.

TABLE 4. *Trends in ratings of school curriculums in drug use prevention (percentages)*

Asked only of those having drug education courses or lectures	High school seniors in the class of: 1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
2E16. Would you say that the information about drugs that you received in school classes or programs has . . .													
1. Made you less interested in trying drugs	50.5	54.0	51.5	52.4	55.3	58.8	56.9	54.7	54.1	55.6	57.2	54.9	58.9
2. Not changed your interest in trying drugs	45.6	43.0	45.2	44.0	41.9	38.5	40.3	42.5	43.3	41.6	40.0	42.8	38.8
3. Made you more interested in trying drugs	4.0	3.0	3.3	3.6	2.9	2.7	2.8	2.8	2.5	2.8	2.9	2.3	2.3
N=	1,973	2,004	2,245	2,006	1,853	2,163	2,022	1,921	1,865	1,953	1,868	2,010	2,110
2E18. Overall, how valuable were the experiences to you?													
1. Little or no value	18.1	19.1	18.0	18.3	16.2	15.4	15.9	18.5	17.8	17.3	17.1	17.9	16.8
2. Some value	45.7	42.6	45.7	44.9	45.2	43.7	44.3	43.0	43.5	43.8	43.8	40.8	38.7
3. Considerable value	24.7	24.6	21.6	22.9	23.6	25.0	23.9	23.7	23.3	24.8	25.5	23.5	25.8
4. Great value	11.4	13.7	14.7	13.9	15.0	15.9	15.9	14.9	15.4	14.0	13.5	17.8	18.7
N=	1,985	1,989	2,237	1,990	1,829	2,159	1,999	1,907	1,857	1,939	1,854	1,991	2,100

SOURCE: Johnston et al., "Monitoring the Future" study.

TABLE 5. *Trends in exposure to antidrug commercials in the media (percentages)*

	High school seniors in the class of:	
The next question asks about antidrug commercials or "spots" that are intended to discourage drug use.	1987	1988
4E11. In recent months, about how often have you seen such antidrug commercials on TV or heard them on the radio?		
1. Not at all	5.7	6.2
2. Less than once a week	6.9	6.2
3. 1-3 times per month	22.1	20.5
4. 1-3 times per week	29.3	31.6
5. Daily or almost daily	25.8	25.4
6. More than once a day	10.2	10.1
	N=2,726	N=2,671

SOURCE: Johnston et al., "Monitoring the Future" study.

more often in recent months, and roughly three-quarters report that such commercials have made them less likely to use drugs. One of the reasons may be that the majority do not see these commercials as having significantly exaggerated the dangers of drugs—in other words, the campaign has retained credibility with the intended audience.

The changes observed in perceived risk, disapproval, and peer norms among these young people with regard to marijuana and cocaine (the two drugs of primary emphasis in the advertisements) are consistent with their positive judgments about the impact of these ads on their own behavior and attitudes. That is not proof-positive that they are having their intended effect, but it certainly is supportive of that interpretation. However, it is interesting to note that the seniors judge the commercials to have less impact on all youngsters their age than the impact they report on themselves in the aggregate.

Presumably, such program evaluations in the large epidemiological surveys could be made considerably more program-specific by getting more detail on exactly which programs the populations were exposed to in school. However, the difficulty of getting accurate information of this sort should not be underestimated.

TABLE 6. *Trends in ratings of antidrug commercials in the media (percentages)*

	High school seniors in the class of:	
4E12a. To what extent do you think such commercials ha. , made people your age less favorable toward drugs?	1987	1988
1. Not at all	22.3	21.0
2. To a little extent	32.8	30.3
3. To some extent	34.3	37.2
4. To a great extent	6.6	7.7
5. To a very great extent	4.0	3.9
	N=2,724	N=2,707
4E12b. To what extent do you think such commercials have made you less favorable toward drugs?		
1. Not all all	25.5	20.5
2. To a little extent	19.9	19.6
3. To some extent	24.6	26.7
4. To a great extent	13.3	14.6
5. To a very great extent	16.5	18.6
	N=2,689	N=2,688
4E12c. To what extent do you think such commercials have made you less likely to use drugs?		
1. Not at all	27.5	23.9
2. To a little extent	17.8	17.0
3. To some extent	21.8	23.7
4. To a great extent	12.5	12.5
5. To a very great extent	20.4	22.9
	N=2,681	N=2,680
4E12d. To what extent do you think such commercials have overstated the dangers or risks of drug use?		
1. Not at all	48.8	49.4
2. To a little extent	16.4	15.5
3. To some extent	18.6	16.7
4. To a great extent	7.4	8.0
5. To a very great extent	8.8	10.4
	N=2,693	N=2,687

SOURCE: Johnston et al., "Monitoring the Future" study.

ASSESSING THE COMBINED EFFECTS OF ALL SOCIETAL FORCES ON DRUG USE

Epidemiological studies, such as the two series of national surveys, provide outcome data on the aggregate impact of all the forces in society that influence drug use—whether they are *labeled* as prevention programs, whether they are *intended* to prevent or promote drug use, and whether they are *organized programs*. In other words, they tell us something about whether all efforts being made, taken in conjunction with all other historical forces of the times, are leading to an improvement or to a deterioration of the situation. In general, these survey series have told us that, for most drugs, there has been an improvement in drug use since as far back as the early 1980s in the majority population that these studies cover. Cocaine was among the last drugs to begin to show such improvement in prevalence rates, which began by 1987 and for "crack" cocaine in 1988 (Johnston et al. 1989). This suggests that "we must be doing something right"—perhaps all of our different prevention efforts in combination (such as political exhortations, media efforts, school prevention programs, individual and collective efforts by parents) are having an effect. One could argue that, in the aggregate, these efforts have considerably more effect than the sum of what can be attributed to each individually because they tend to reinforce or resonate on one another.

We must be careful to recognize that not all efforts intended to reduce drug use occur as a result of formal prevention programs. For example, not all programs that are effective in reducing drug use are labeled as prevention programs ("deterrence programs," for example, might be the label used for criminal or civil law enforcement interventions), and not all efforts intended to reduce drug use are systematic *programs* (much occurs as one-to-one social influence attempts, social modeling, etc.). Finally, we must remember that not all forces that reduce drug use develop because of someone's *intent* to reduce other people's drug use. A useful definition of "drug use prevention activities" in their broadest form might be this: Drug prevention activities encompass any actions, or programs of action, undertaken by individuals or other social entities for the purpose of preventing the onset of, or reducing the use of, one or more drugs by certain individuals or groups in the population.[2]

This definition would include social influence that occurs in one-to-one interaction—for example, between parent and child or between sibling and child—and intentional social modeling. It would not include vicarious learning (e.g., from the experiences of Len Bias and Don Rogers or from others in an individual's immediate environment) even though these also may be important determinants. It also would include some programs intended to reduce drug use not necessarily encompassed under what most of us think of as "prevention

programs"—for example, changes in laws, enforcement, adjudication, and penalties. These efforts are not initiated to win the hearts and minds of the people, as are most traditional prevention programs, but they are directed at changing the punitive contingencies attached to drug-using behaviors. They do not work through a change in knowledge, attitude, beliefs, or norms. They operate by changing the reinforcement schedule offered by the formal system and by letting specific consequences be known. Whether such changes have been or will be effective in the domestic population is highly conjectural. The evidence is strong, however, that within the more "total environment" of military service, such changes can make and have made a difference in drug-using behavior (Bray et al. 1988).

Finally, not all historical forces that reduce drug use are *intended* to reduce drug use—in fact, many of them simply occur. To take one important example, it has been argued that the Vietnam War did much to stimulate drug use and that its passing has had the effect of removing an important catalyst to use—symbolic expression (Johnston 1973; Johnston et al. 1987). Also, the passing of the baby-boom generation into young adulthood lowered the absolute number of adolescents who might use drugs. The recession of the early 1980s, which in combination with the baby boom created a real shortage of entry-level jobs, may well have made young people more career conscious, more concerned about academic achievement, and less likely to use drugs. The point is that historical shifts in factors as broad as war, recession, and the age composition of the population may be major determinants of changes in drug use. It appears that shifts of all three types conspired to help bring about the massive epidemic of illicit drug use in the 1960s and 1970s and may now be contributing substantially to the declines in use during the 1980s (Johnston, in press). Thus, it would be faulty to credit all improvement to society's "drug prevention activities," even as broadly defined above.

But there is also evidence from the Monitoring the Future study that certain changes that probably *are* attributable in part to drug prevention efforts are occurring. However, these efforts are focused on the dangers judged to be associated with the use of these drugs. Here, beliefs associated with the use of specific drugs change and, because these beliefs have been the target of intended change in most school-based curriculums and certainly in the major antidrug media campaigns, it seems likely that some of that change is due to informative efforts. Figures 5 and 6 show the changes for high school seniors in perceived risk for marijuana and cocaine and the concurrent changes in use of these drugs. The shifts in perceived risk have been dramatic, and various pieces of evidence presented elsewhere are convincingly influential in reducing use (Bachman et al. 1988; Johnston 1982, 1985, in press; Johnston et al. 1980, 1989). It seems clear from figures 5 and 6 that a shift in availability was not the cause of the downturn for either drug.

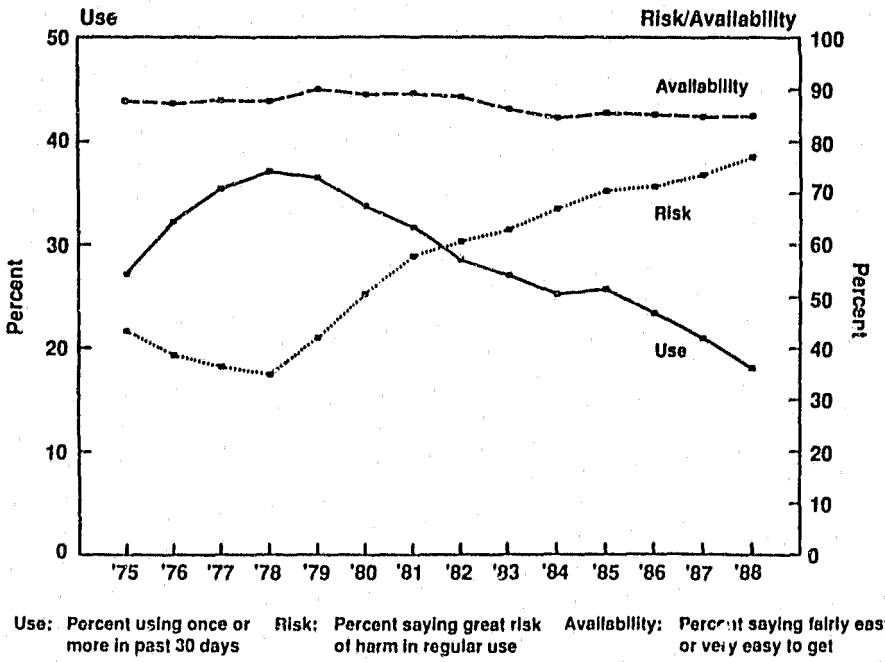

FIGURE 5. *Trends for marijuana in perceived availability, perceived risk of regular use, and use in past 30 days, 1975-88—high school seniors*

SOURCE: Johnston et al. 1989.

There are additional forces to planned prevention programs that have helped to create and disseminate evidence about the risks of these two drugs. These forces have included the accumulation of clinical, laboratory, and epidemiological evidence on the effects of these drugs as well as the dissemination of that evidence by the media. Also, many young people have observed firsthand the experiences of friends and acquaintances who were users, which provided the opportunity for vicarious learning and which seems particularly true for marijuana, although planned influence attempts by schools, families, and the media undoubtedly were important conduits. Certainly, high exposure to the media campaigns by young people and their favorable credibility and impact ratings suggest that they also have played an important role.

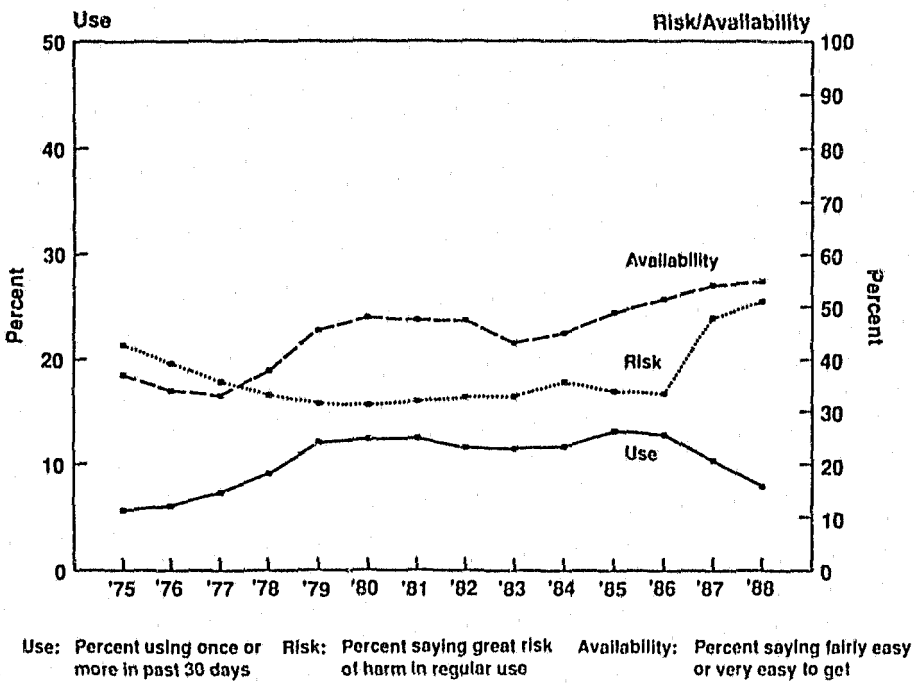

FIGURE 6. *Trends for cocaine in perceived availability, perceived risk of trying, and use in past year, 1975-88—high school seniors*

SOURCE: Johnston et al. 1989.

In summary, epidemiological survey data provide useful prevention-related information regarding the combined influences of intended programmatic, unintended programmatic, and other historical forces. Sharp changes in slope at a given point that correspond to new programmatic interventions provides evidence of a quasi-experimental nature that the program had an effect or that "something is right." Again, this kind of information is certainly no substitute for carefully designed evaluations of programmatic interventions, but it provides a form of inferential evidence about society's collective efforts.

CONCLUSION

To conclude this discussion of the various ways in which epidemiological (including etiologic) research can be and has been helpful to the development

of prevention programs and program evaluations, it should be noted that this has not been an exhaustive review. For example, providing good instrumentation and good field procedures for evaluation studies have not been mentioned.

In addition, there can be no doubt that prevention research can make and has made several constructive contributions to epidemiology—particularly related to theory. There is nothing like changing an individual characteristic or an environmental factor to see if it is really a determinant of use. Further, prevention specialists often do much of the critical *qualitative field* research needed to expand perspectives and theories about the epidemiology and etiology of drug use, but these could constitute the laws of a separate chapter.

Clearly, the field of drug epidemiology has influenced and should continue to influence the field of drug abuse prevention in several ways. Because of the dynamic nature of all of the phenomena under study, that influence undoubtedly will continue in the coming decade. Indeed, the work of these two fields should be more closely integrated to advance the work of each of them at a time when their contributions are so important to containing the Nation's drug use epidemics.

NOTES

1. We have shown that for most drugs the similarities in use across different levels of urbanicity are more impressive than the differences.
2. Drug *use* prevention is distinguishable from drug *abuse* prevention. Programs or actions included in the latter area may have more limited objectives of preventing or reducing the adverse *consequences* from the use of drugs, without necessarily preventing or reducing use. (For instance, a goal might be to reduce alcohol-impaired driving without reducing occasions of alcohol impairment by establishing norms about not driving when alcohol impaired.) Or, if a type of use results in most of the adverse consequences, the objective may be to avoid only that type of use. (The call for "responsible use" would fit into this category.)

REFERENCES

Bachman, J.G.; Johnston, L.D.; and O'Malley, P.M. Smoking, drinking, and drug use among American high school students: Correlates and trends, 1975-1979. *Am J Public Health* 71:59-69, 1981.

Bachman, J.G.; Johnston, L.D.; O'Malley, P.M.; and Humphrey, R.H. Explaining the recent decline in marijuana use: Differentiating the effects of

perceived risks, disapproval, and general lifestyle factors. *J Health Soc Behav* 29:92-112, 1988.

Bachman, J.G.; O'Malley, P.M.; and Johnston, L.D. *Change and Consistency in the Correlates of Drug Use Among High School Seniors: 1976-1986.* Monitoring the Future Occasional Paper No. 21. Ann Arbor, MI: Institute for Social Research, 1986.

Bray, R.M.; Marsden, M.E.; Guess, L.L.; Wheeless, S.C.; Iannacchione, V G.; and Keesling, S.R. *Highlights of the 1988 World Survey of Substance Abuse and Health Behaviors Among Military Personnel.* Research Triangle Park, NC: Research Triangle Institute, 1988.

Jessor, R., and Jessor, S.L. *Problem Behavior and Psychological Development: A Longitudinal Study of Youth.* New York: Academic Press, 1977.

Johnston, L.D. *Drugs and American Youth.* Ann Arbor, MI: Institute for Social Research, 1973.

Johnston, L.D. A review and analysis of recent changes in marijuana use by American young people. In: *Marijuana: The National Impact on Education.* New York: American Council on Marijuana, 1982. pp. 8-13.

Johnston, L.D. The etiology and prevention of substance use: What can be learned from recent historical changes? In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 155-177.

Johnston, L.D. Adolescent smoking and the issue of cigarette advertising. Prepared testimony delivered before the House Subcommittee on Health and the Environment in oversight hearings on cigarette advertising and promotion. *Advertising of Tobacco Products.* Serial No. 99-167. Washington, DC: U.S. Govt. Print. Off., 1987. pp. 860-886.

Johnston, L.D. *Reducing Drug Use in America: A Perspective, a Strategy, and Some Promising Approaches.* Monitoring the Future Occasional Paper No. 26. Ann Arbor, MI: Institute for Social Research, 1989.

Johnston, L.D. Toward a theory of drug epidemics. In: Donohew, L.; Sypher, H.; and Bukoski, W., eds. *Persuasive Communication and Drug Abuse Prevention.* Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Johnston, L.D.; Bachman, J.G.; and O'Malley, P.M. *Highlights From Student Drug Use in America, 1975-1980.* DHHS Pub. No. (ADM)81-1066. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1980.

Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. *Lifestyle Orientations in Late Adolescence and Patterns of Substance Abuse.* Monitoring the Future Occasional Paper No. 27. Ann Arbor, MI: Institute for Social Research, 1987.

Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. *Drug Use, Drinking, and Smoking: National Survey Results from High School, College, and Young*

Adult Populations, 1975-1988. DHHS Pub. No. (ADM)89-1638. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1989.

Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. "Monitoring the Future Study: A Continuing Study of the Lifestyles and Values of Youth." National Institute on Drug Abuse. University of Michigan, Institute for Social Research, Ann Arbor, MI.

Kandel, D.B. Stages in adolescent involvement in drug use. *Science* 190:912-914, 1975a.

Kandel, D.B. Reaching the hard-to-teach: Illicit drug use among high school absentees. *Addict Dis* 1:465-480, 1975b.

Kandel, D.B. Convergences in prospective longitudinal surveys of drug use in normal populations. In: Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere Publishing, 1978.

Miller, J.D.; Cisin, I.H.; Gardner-Keaton, H.; Harrell, A.V.; Wirtz, P.W.; Abelson, H.I.; and Fishburne, P.M. *National Survey on Drug Abuse: Main Findings, 1982.* DHHS Pub. No. (ADM)83-1263. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983.

National Commission on Marijuana and Drug Abuse. *Marijuana: A Signal of Misunderstanding* (1st report). Stock No. 5266-0001. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1972.

National Institute on Drug Abuse. *National Household Survey on Drug Abuse: Main Findings 1985.* DHHS Pub. No. (ADM)88-1586. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1988.

O'Malley, P.M.; Bachman, J.G.; and Johnston, L.D. Period, age, and cohort effects on substance abuse among young Americans: A decade of change, 1976-1986. *Am J Public Health* 78:1315-1321, 1988.

Osgood, D.W.; Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. The generality of deviance in late adolescence and early adulthood. *Am Sociol Rev* 53:81-93, 1988.

AUTHOR

Lloyd D. Johnston, Ph.D.
Program Director
Institute for Social Research
University of Michigan
Ann Arbor, MI 48106-1248

140140

Methodological Issues in Drug Use Prevention Research: Theoretical Foundations

Brian R. Flay and John Petraitis

INTRODUCTION

This chapter concerns the theoretical foundations of drug use prevention program development and research. We first briefly discuss the nature of theories in the social sciences and then actual theories of drug use onset and behavior change; finally, we focus on the functions and roles of theory and their methodological applications.

WHAT IS THEORY?

"A theory is a systematically related set of statements, including some law-like generalizations, that is empirically testable" (Rudner 1966, p. 10). Theories or models are abstractions that enable us to represent part of the world by a simpler structure. They simplify decisionmaking and help us predict the future. Most important, they help us generalize from one instance, location, or group to another.

Most social science theories provide us with "orienting statements" (Homans 1967) about social phenomena rather than with strictly mathematical laws that are more common in the basic sciences.[1] These orient us to look for determinants of social phenomena in certain places rather than others. Thus, when considering the determinants of children's drug use, Marxists look for causes in the economic structure of society rather than in individual development; biologists (and many with medical training) may look for genetic determinants; sociologists look to patterns of social interaction and social influences of parents, peers, and advertising; and psychologists look to personality and cognitive areas. All such orienting statements and the theories derived from them may be plausible, and they are not mutually exclusive (Biglan and Lichtenstein 1984). Any given view of reality reflects as much the theoretical perspective or methods of the observer as it does the object being

viewed (Campbell 1969). However, one perspective may be more useful than others in specific applications.

Theories of Drug Use Onset

Researchers have developed numerous theories of drug use. Lettieri and colleagues (1984) analyzed 43 of them. Murray and Perry (1985) and Newcomb and Bentler (1988) each provide analysis of smaller sets of the major theories. With few notable exceptions, most of these theories were derived from narrow disciplinary perspectives and on the basis of cross-sectional correlates of drug use.

Predictors of Drug Use. There is agreement on the major predictors of drug use, and reviews of the correlates of drug use are numerous (e.g., Braucht et al. 1973; Flay et al. 1983; Gorsuch and Butler 1976; Hawkins et al. 1985; Huba and Bentler 1980; Jessor 1979; Jessor and Jessor 1977; Johnston et al. 1982; Kandel 1978a, 1978b, 1980, 1982; Kandel et al. 1978; Lettieri and Ludford 1981; Miller et al. 1983; Murray and Perry 1985; Sadava 1987; Smith and Fogg 1978; Wingard et al. 1979, 1980). Hawkins and colleagues (1985) provide by far the most thorough and comprehensive review. However, one is struck by the large number of correlates in the absence of any theoretical framework (Shore 1985). First, many correlates are found only in cross-sectional studies; second, there is no information about the relationships among the correlates; and third, investigators use different labels with different orientations for the same phenomenon or construct or the use of the same label for different constructs or phenomena.

To build a parsimonious theoretical framework, we consider the predictors found in 24 prospective studies covering the childhood through young adult years (table 1). There is reasonable consistency among studies in the domains of variables found to predict drug use prospectively. Figure 1 shows five classes of variables that we believe encompass the most important predictors of drug use confirmed repeatedly in prospective studies. They are also common to the most developed and integrated theories of drug use behavior.

Knowledge, Attitude, and Behavior (KAB) Change. Starting from the right of figure 1, and the most proximal to actual drug use, are the intrapersonal cognitive, affective, and conative variables, or KAB. These include knowledge of physiological and social consequences of use; personal beliefs (expectancies, perceived risk, susceptibility) regarding consequences; general values (e.g., toward health, independence) and specific evaluations of these consequences; attitudes toward drug use and related issues; behavioral intentions; trial behavior; stages of behavior (e.g., alcohol, tobacco, marijuana,

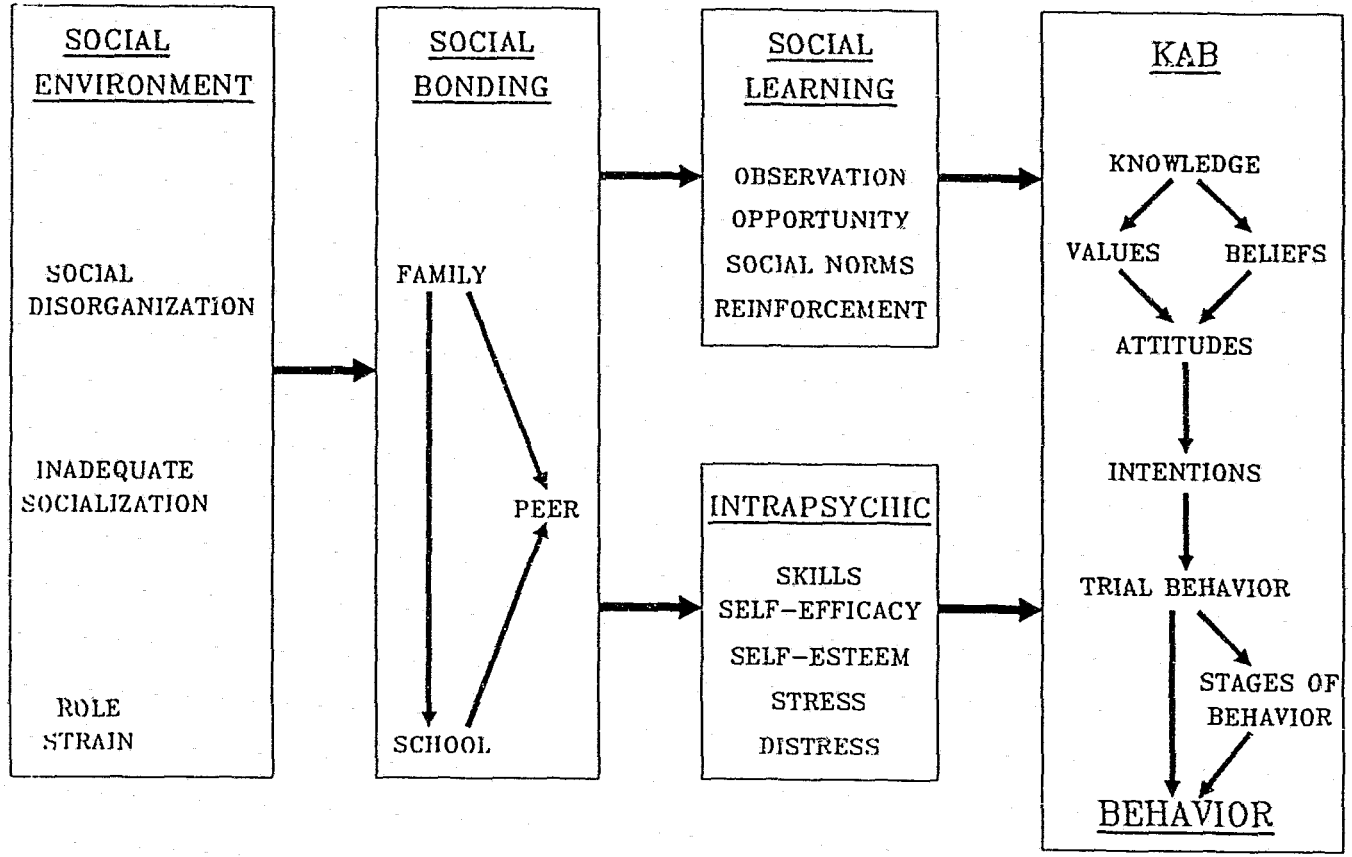

FIGURE 1. *Five domains of determinants of drug use/abuse*

TABLE 1. *Characteristics of longitudinal studies of teenage alcohol and other drug use*

Data Set	Year	Place	Age	Time	N	Percent Attr	Alcohol DV	Alcohol Baseline	Alcohol Endline	Marijuana DV	Marijuana Baseline	Marijuana Endline	Harder Drugs	Analytic Method	Theory Tested
Childhood predictors of teenage alcohol and other drug use															
1 Baumrind	1968	?	4	10yr	134	22	6	0	36	6	0	64	Y	MR	Baumrind
2 Block	1969	Berkeley	3	11yr	130	19	2	0	65	6	0	51	Y	r	–
3 Brook	?	NY City	5-10	8yr	503	39	2	0	75	2	0	33	Y	r,MR	Brook
4 Kellam	1966	Chicago	6	10yr	705	25	6	0	73	6	0	38	N	S	–
5 Vicary	1956	NY City	1	19yr	133	0	5	0	74	5	0	47	N	MR	–
Early teenage predictors of later teenage alcohol and other drug use															
1 Bailey	1985	SE USA	12-14	18mo	7,562	54	–	–	–	2	13	25	–	LR	Elliott
2 Brook	1986	USA+	14-15	2yr	932	24	5	55	80	–	–	–	N	r	–
3 Christian	1989	Detroit	12-13	1yr	871	23	6	52	68	–	–	–	N	r,MR	–
4 Elliott	1976	USA+	11-17	3yr	1,725	6	Y	?	?	Y	?	–	Y	P	Elliott
5 Friedman	–	Philadelphia	14-16	17mo	598	?	9	?	?	9	?	–	Y	CL	–
6 Galambos	1982	Berlin	11-15	2yr	622	38	5	?	?	–	–	–	N	P	Kaplan
7 Jessor	1969	Colorado	12-14	4yr	589	18	3	49	74	3	16	38	N	F	Jessor
8 Kandel	1971	NY State	15-16	5mo	8,206	34	8	17	21	8	19	23	Y	MR,L	Socialization
9 Kaplan	1971	Houston	12	2yr	7,618	58	W	9	21	M	2	12	Y	r,P	Kaplan
10 Smith	1969	Boston	12-13	5yr	2,249	?	6	?	?	6	8	33	Y	T,F	–
11 Teichman	1982	Israel	15-18	1yr	1,900	24	2	?	?	2	?	–	Y	F	Kandel
12 UCLA	1976	Los Angeles	12-14	5yr	1,634	48	5	?	?	5	?	–	Y	r,L,CC	Domain
13 Windel	1979	USA+	14-15	4yr	2,411	5	6	46	?	6	23	–	–	MR	Jessor
Predictors of young adult use and abuse															
1 Campbell	1961	N. Carolina	17	1yr	1,225	1	2	17	38	–	–	–	N	None	–
2 Jessor	1969	Colorado	16-17	10yr	865	32	C	81	92	–	–	–	N	r,t	Jessor
3 Johnston	1966	USA+	15	5yr	2,213	27	5	81	89	5	21	34	Y	CL	–
4 Kandel	1971	NY State	15-16	9yr	8,206	44	8	50+	99	Y	14	43	Y	L	Socialization
5 Kaplan	1971	Houston	12	10yr	7,618	61	–	–	–	2	8	65	N	r,P	Kaplan
6 Pulkkinen	1968	Finland	8	12yr	369	63	C	0	90	2	0	4	N	F	Pulkkinen
7 Schlegel	1974	Ontario	15-18	6yr	1,752	48	6	80	90	6	30	50	N	MR	Jessor, Fishbein
8 UCLA	1976	Los Angeles	12-14	8yr	1,634	61	5	?	?	5	?	?	Y	r,L	–

NOTE: Question marks denote that information was not reported or could not be derived from published sources.

KEY: Year = year that data collection began
Age = age level at beginning of study
Time = from beginning of study to followup
N = sample size at beginning of study
Percent Attr = percent of initial subjects with data at the final wave
DV = dependent variable: 2 = 2 point scale of lifetime use (0 = never used)
3 = 3 point scale of lifetime use (0 = never used)
5 = 5 point scale of lifetime use (0 = never used)
6 = 6 point scale of lifetime use (0 = never used)
8 = 8 point scale of lifetime use (0 = never used)
9 = 9 point scale of lifetime use (0 = never used)
CU = current use (0 = current abstinence)
W = used during past week (0 = not in past week)
M = used during past month (0 = not in past month)
Y = used during past year (0 = not in past year)
Baseline = estimated percent of users at time of first measurement
Endline = estimated percent of users at time of last measurement
Harder Drugs = whether predictors of harder drugs (e.g., cocaine) were reported
Analytic Method: t = t-tests r = correlation CL = cross-lagged panel analysis
F = ANOVA MR = multiple regression L = Lazarsfeld 4-fold table
M = MANOVA CC = canonical correlations LI = LISREL/structural analysis
P = path analysis S = survival analysis LR = logistic regression
+ = national probability sample

harder drugs, etc. (Kandel 1975); and established adult behavior patterns. Theories of how these variables relate to each other abound, and we have attempted some integration of these and other theories and variables in figure 2 (Flay 1981; Flay et al. 1983).

Social Learning Theory. The second set of variables for which relationships to behavior are well established are the social learning variables of opportunities for observation and modeling of the behavior; opportunities to use (or availability); social normative beliefs, including collective ignorance of norms; and social reinforcement (positive and negative). Bandura (1977, 1986) and Akers (1977) have developed the relations among these variables from the psychological and sociological perspectives, respectively. Any particular behavior is more likely to occur when it is differentially reinforced and is seen as desirable by important others. Fishbein and Ajzen (1975) and Ajzen and Fishbein (1980) have incorporated social normative beliefs into their theory of reasoned action for the prediction of behavioral intentions. Several groups (Akers and Cochran 1983; Akers et al. 1979; Elliott et al. 1985; Hawkins et al. 1986; Hawkins and Weis 1985) also have applied Akers' version of social learning theory to predict and explain delinquency behavior, including drug use (see below).

Social Environment. Sociologists and others have established that social environment variables most distal from behavior (see the left of figure 1) predict drug use. Thus, the structure of the economic, legal, social, and educational systems of a society are determinants of behavior. In particular, role strain (Merton 1957) and social disorganization or breakdown may lead to inadequate socialization that in turn alters the social bonding and social learning variables (e.g., observation, opportunities). These may then lead to increased drug use among, for example, the disadvantaged (Wilson 1987).

Social Bonding. Sociologists believe that the mechanisms through which social organization affects drug use concern social bonding. Thus, conventional bonds with family, peers, school, and other community groups are important. For example, researchers have shown that breakdown of family bonding leads to increased probability of bonding with delinquent peers as shown in figure 3 (Elliott et al. 1985).

Intrapsychic Variables. Both sociologists and psychologists have suggested that "intrapsychic" variables might complete the link between social bonding and KAB variables. Sociologists suggest that poor family bonding leads to stress (inability to cope, rebelliousness, risk-taking) and distress (withdrawal, self-derogation, depression). Several researchers have shown a link between social stress or distress variables and substance use (Shiffman and Wills 1985;

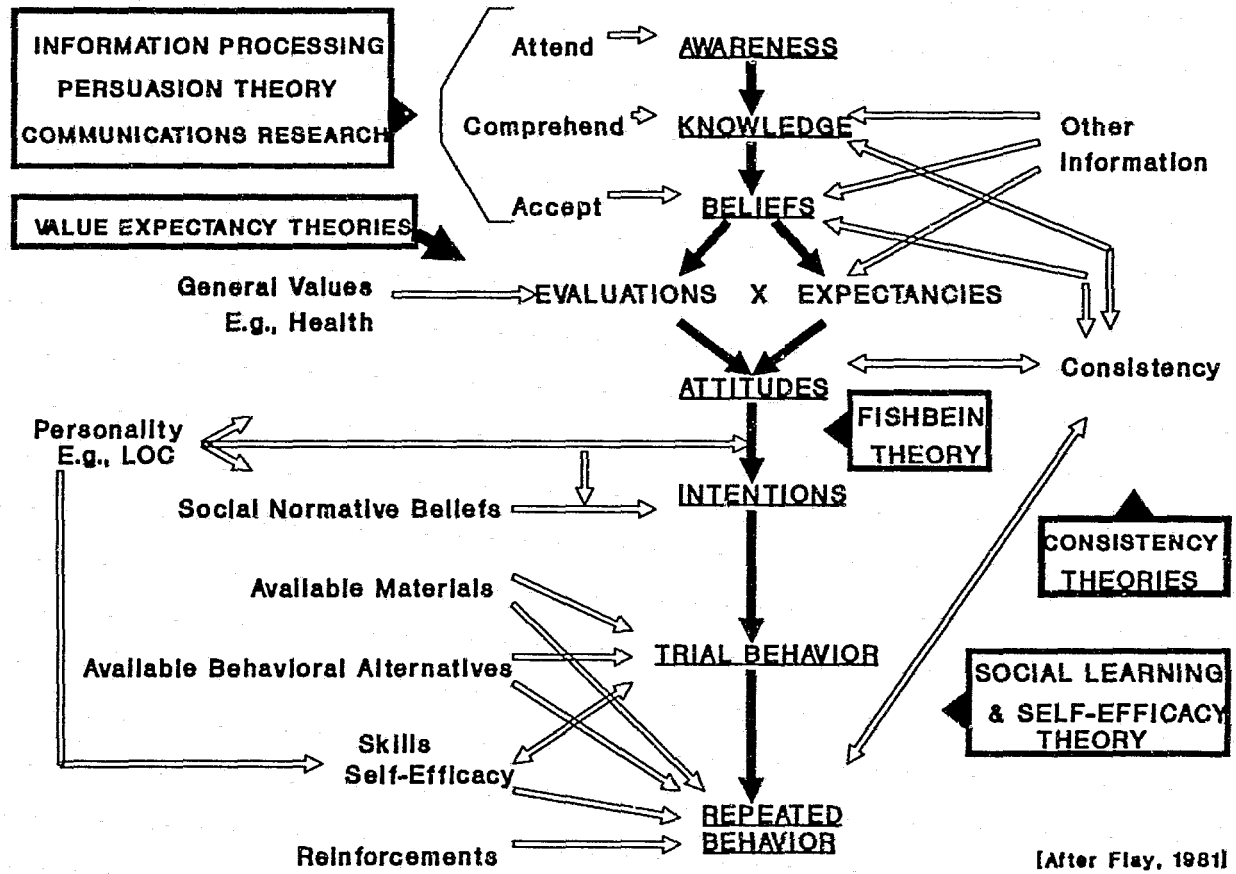

FIGURE 2. *Flay's integrated psychological model*

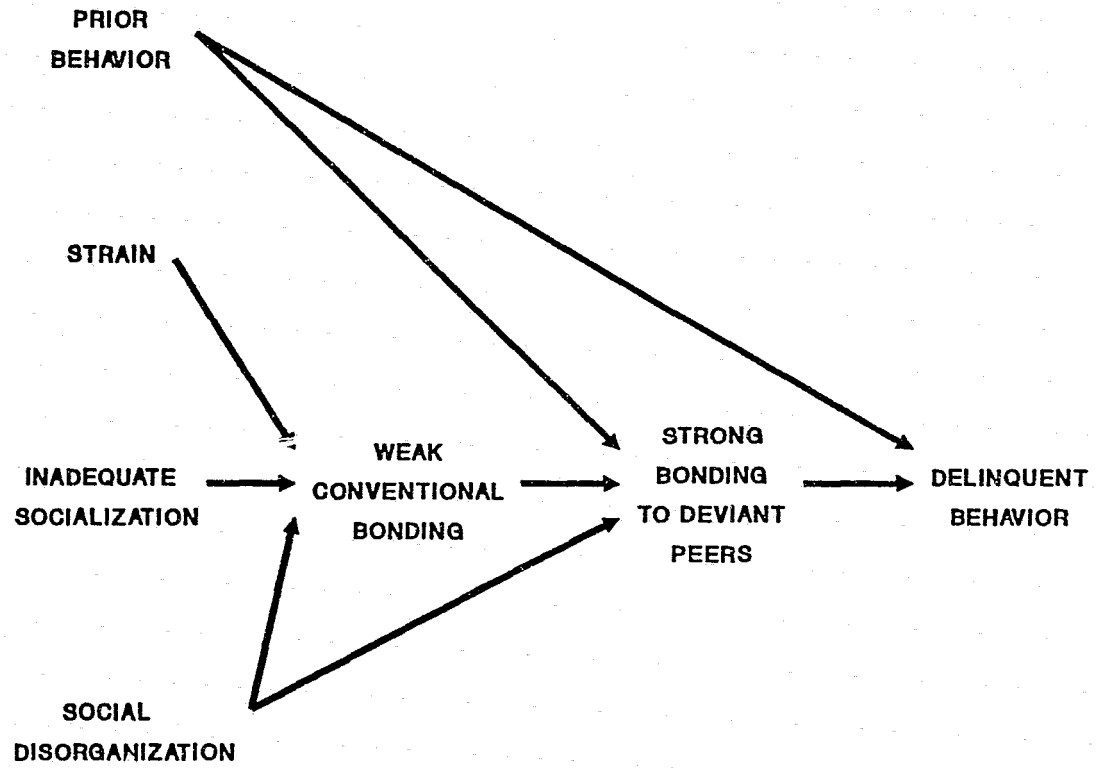

FIGURE 3. *The Elliott et al. integrated sociological model*

SOURCE: Elliot et al. 1985, p. 146.

Shontz and Spotts 1986; Winick 1974, 1986) as early as grade 1 (Kellam et al. 1982). On the other hand, strong family and other conventional bonding can lead to the development of positive social skills and competencies, strong self-efficacy regarding these, and high self-esteem. Psychologists have suggested that personality factors (e.g., locus of control) affect one's ability to cope with social situations and one's desire for and response to drug use (Kaplan 1975; Kaplan et al. 1982, 1984).

Interactional theorists emphasize interactions between personal and environmental variables in addition to independent effects (Sadava 1987; Sadava and Forsyth 1977). Protective combinations may be found by examining interactions. For example, the increased risk of drug use due to childhood depression may be decreased or eliminated by having a nonworking parent or it may be increased by latchkey status (Richardson et al. 1989). Brook and colleagues (1984, 1985) found that adolescents with poor psychological adjustment and lack of goal orientation were less at risk if their mothers were psychologically stable.

Summary. The above-mentioned five classes of variables seem to encompass all variables included in other broad theories. For example, the Jessor and Jessor (1977) personality, perceived social environment, and behavior variables are all accounted for by the above domains, as are most of the domains in the University of California-Los Angeles domain theory (Huba et al. 1984; Huba and Bentler 1980, 1982; Newcomb et al. 1983). We hope that this particular integration of the predictors of drug use provides more orienting statements about macrolevel relationships and processes (both psychological and sociological) than any other single theory or model. However, we are still left with a certain uneasiness; that is, we may have suggested greater agreement between different investigators' theoretical constructs and points of view than they would accept. On the other hand, some such integration is necessary if we ever are to make progress in understanding the onset of drug use and abuse[2] (Akers 1989). Specific theories and integrations within domains should continue to guide microlevel relationships.

FUNCTIONS OF THEORY

What is the value of all this theory, integrated or not? We now consider 12 functions or roles of theory in the development and testing of drug use prevention programs in three sections: roles in program development, roles in program evaluation, and larger scientific functions.

Roles in Drug Use Prevention Program Development

Link Theoretical Elements and Program Components. Theories of drug use provide us with guidelines or orienting statements for program development. They show us the intervening variables that need to be targeted and suggest the type of curriculum components that might be effective. For example, a focus on the KAB domain suggests informational and values clarification approaches. A focus on social learning domains suggests clarification and correction of social influences and training in social skills. The intrapsychic domains suggest self-esteem enhancement and stress-coping approaches. The social bonding domain suggests family communication patterns and the use of peer leader/models. The social environment domain suggests improved socialization and large changes in society.

Reach Consensus Regarding Magnitude of Program Effects. Formal social science theories should inform us of the magnitude of program effects that we can realistically expect from any particular program. Careful consideration of the probabilistic and stochastic nature of most links in social science theories would lead to more conservative estimates of preventive effects of any particular approach than those that we have made in the past. For example, we should not expect information to play a large role when knowledge and beliefs can be influenced by so many other sources and when the translation of knowledge into behavior depends on so many other factors. As another example, we should never have expected that teaching social skills in a one-shot program could, by itself, lead to large and lasting changes. Certainly, the more distal the intervening variable targeted by an activity, the smaller can be the expected effect on behavior. Similarly, the greater the number of intervening variables left unaddressed by an intervention, the smaller its effects on behavior should be.

Suggest Need for Comprehensive Programs. As suggested above, any one microlevel theory describes only a portion of the total phenomenon of drug use; thus, integration would appear to be necessary to obtain a more complete view. We also saw that many of the theories overlap considerably and that this overlap of various theories also suggests a need for integration (figure 2). They all demonstrate the need to consider multiple theories and operationalizations to understand drug use to develop and evaluate prevention programs. Reliance on any one theory or discipline is inappropriate and inadequate (Evans 1988). They all suggest that an effective prevention program must be comprehensive, which probably also means delivery over an extended period of years rather than months, weeks, or days.

Roles in Prevention Program Evaluation

Link Program Components and Intervening Variables. All program effects are achieved through a set of short-term processes involving intervening variables. Each component of a program is designed to affect a certain intervening variable or set of variables. Theory enables us to specify these linkages, measure the appropriate variables, and conduct the appropriate analyses of process (Dwyer 1983; Judd and Kenney 1981). For example, Project STAR (Students Taught Awareness and Resistance) includes components to correct knowledge of consequences of drug use, improve communication with friends, and decrease intentions to use drugs in the future. We have been able to demonstrate that these variables did indeed change as hypothesized[3] (MacKinnon et al., unpublished manuscript). In another study, the Television, School, and Family Project, we found that social influences information changed social influences knowledge/beliefs and not knowledge of consequences and vice versa (Sussman et al. 1989).

Inform Program Implementation/Dissemination. No program is effective unless implemented appropriately so as to reach the intended audience, hold their attention, be credible to them, and teach them something. Theory can inform us about program delivery. How a program is delivered, by whom, and under what conditions (settings) all intervene between the program and its effects. For example, whether a curriculum is delivered by a teacher or a peer (or both) might determine how well students accept and learn from it (Brannon et al., in press). Fidelity of curriculum delivery will determine its effects on theoretically relevant intervening variables (Sobol et al. 1989).

Inform External Validity. Program theory should inform us as to the external validity or generalizability of program effects by target audience and social environment characteristics.

Target audience characteristics and interactions. Target audience characteristics may interact with program approaches or components (Lipsey et al. 1985), and theory should be able to inform us of these potentials. For example, social influence prevention programs may be effective for those students at risk of becoming drug users because of social influences but not those at risk because they are rebellious.

Inform social environment characteristics and interactions. Theory also can inform us of possible interactions between the social environment and program approaches or components (Lipsey et al. 1985). For example, programs might be more effective in schools with articulated policies supporting a drug-free environment than in others (Pentz et al., in press), and we might

expect some prevention programs to be effective in middle-class suburbs but not in disadvantaged neighborhoods.

Inform Construct Validity. All program treatments must operate through a set of theoretically derived intervening processes to produce the desired effects.

Clarify cause-effect relationships and intervening variables. Theory can help us design programs and measures to address cause-effect relationships and many intervening variables. For example, in Project STAR, we not only demonstrated program effects on intervening variables but also showed that these changes mediated behavior change (MacKinnon et al., unpublished manuscript).

Suggest short-term versus long-term effects. All long-term goals must be achieved through a set of short-term processes or effects. For example, one goal of most prevention programs is to lower the prevalence of drug use by the end of high school or after. This might be achieved by reducing the intentions to use drugs or the probability of trying drugs earlier. Theory helps us specify these variables and their relationships and the timeframe over which it is reasonable to expect certain levels of effects.

Suggest unintended effects. The same intervening processes needed to produce desired effects also may produce other, perhaps unintended and undesirable, effects. For example, increasing knowledge of the consequences of drug use might lead to a lower desire to use them and less use, or it might lead to an increased desire to try them (Goodstadt 1980). Decreasing cigarette use among male adolescents may be substituted with increased use of smokeless tobacco. Theory should help us foresee potential unintended program effects. Alternatively, detection of unintended effects can enlarge our knowledge base about drug use and its prevention (Chen and Rossi 1980).

Inform Measurement. Theory can inform us what to measure when testing or evaluating a prevention intervention. The important classes of variables are expected outcomes, intervening variables, implementation processes, program content, audience characteristics, environmental and setting characteristics (e.g., school characteristics), and unintended effects.

Help Explain Effects of Nontheoretically Derived Programs. Sometimes practitioners with little or no theoretical background may design effective prevention programs. Application of theory during a formal evaluation can help explain the effects of such programs, and once such effects are theoretically understood, they may be improved or built on.

Large Scientific Roles

In addition to specific roles of theory in program development and evaluation, theories of any behavior and its alteration allow us to (1) discriminate between program and theory failure, (2) contribute to social science knowledge, and (3) contribute to research efficiency.

Discriminate Program Failure and Theory Failure. Failure to find program effects can be due to wrong theory, poor translation of theory into program, poor program implementation, or poor evaluation design (Bickman 1987; Suchman 1967; Weiss 1972). Theory failure has occurred if intervening variables have changed as hypothesized but behavior has not. Program failure has occurred if expected changes did not occur. This determination is possible only if the evaluation and program have a strong theoretical basis.

Contribute to Social Science Knowledge. Tests of prevention programs can be an important source of social science data (Chen and Rossi 1983). Use of theoretically meaningful program process and outcome variables, that is, variables with high construct validity (Cook and Campbell 1979), can lead to important contributions to social science (Bickman 1987). Tests of prevention programs can be just like basic research in that we attempt to understand the relationships between program variables and outcomes. There are many gaps in our understanding of drug use; good theory-based research and evaluation can help fill these gaps.

Improve Long-Term Research Efficiency. Better understanding of why each evaluated program was or was not effective will lead to more efficient research efforts in the long term.

IMPLICATIONS FOR METHODS

We discuss the implications of theory for methods of research in four major groups: implementation quality, external validity, construct validity, and special method-theory relationships. The major implications for each of the first three areas concerns measurement and analysis. In all future research, we need to construct indicators of implementation quality, external validity, and construct validity and link variations in them with ultimate program effects.

Implementation Quality

Theories of drug use onset and behavior change can inform program delivery/ implementation as well as program content; however, a specific theory of program implementation also might be helpful (Chen and Rossi 1983). At the

simplest level, ultimate program effects depend not only on program content but also on mode and quality of delivery and on attention and learning by the audience. Determinants of quality of delivery and students' attention include such factors as (1) the nature of the program and its content; (2) social, political, and financial support at the school district and community levels; (3) teacher (and peer) training; and (4) acceptance of the program by teachers, parents, and students (figure 4).

Thus, in evaluating drug prevention programs, we should measure program content, implementation methods and integrity (Sechrest et al. 1979), and program acceptance by students, parents, teachers, other officials, and other community groups. We then need to conduct analyses to establish links between each of these and program effects.

External Validity

We know very little from studies conducted to date about the external validity or generalizability of their findings. Generalizability concerns the transferability of an effective program—for whom is it effective and under what conditions of implementation/dissemination? Most rigorous studies of recent smoking prevention programs, for example, have been conducted on white middle-class populations (Flay 1985). We still do not know for sure whether they are effective for various socioeconomic and ethnic groups.

We know relatively little about the students for whom the psychosocial approaches are most effective. Most studies have not performed separate analyses by sex, grade, prior experience with substances, or other characteristics of the study participants. Where such analyses have been done, differences have sometimes been found; for example, results from some studies of the social influences approach to smoking prevention suggest that males and females are equally influenced by a teacher-led program but that they may be differentially influenced by a peer-led program (Flay 1985).

Another area that past research has not yet addressed sufficiently concerns broader issues of program dissemination. Once we have an efficacious program, how will it be disseminated? Should regular teachers be trained? If so, how? Would some other group, such as school nurses or health agency volunteers, be more effective? What is the potential role of the media? Will those programs found to be most efficacious under research conditions also be found most effective under real-world conditions (Flay 1986)? All such questions remain for further research to answer.

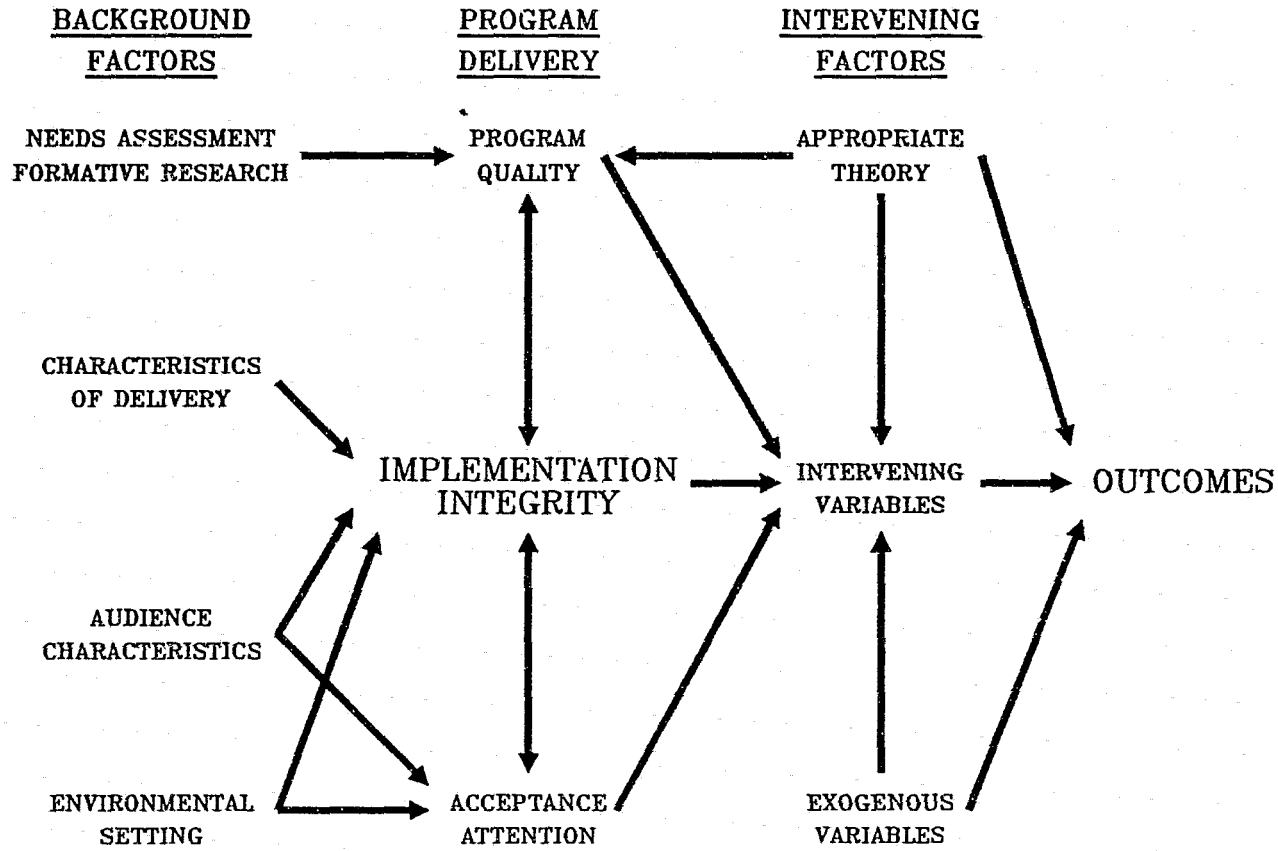

FIGURE 4. *Model of program delivery*

Construct Validity

Construct validity of the treatment concerns questions of whether the various components of a program have the immediate effects expected of them and whether any immediate effects on presumed mediating variables are related to subsequent behavior.[4] Each component of a program is designed to produce a particular effect, and it is the combination of all those effects that should prevent drug use. Few past studies reported program effects on presumed mediating variables, and even fewer attempted to link any such changes to subsequent drug use behavior. The investigations of the more general life/social skills approaches have been more diligent at including assessments of presumed mediating variables (McCaul and Glasgow 1985). An analysis by Glasgow and McCaul (1985), however, demonstrates great inconsistency across studies in those mediating variables affected—even by the same or very similar programs tested by the same researchers—and no attempts to link changes in mediating variables to behavior change.

In addition to measuring and describing their program and its implementation in detail, future researchers will need to assess program effects on presumed mediating variables (e.g., attitudes, intentions, resistance skills) and attempt to link changes in presumed mediators with changes in subsequent smoking behavior. Such research also will enhance our knowledge about the process of becoming a drug user, which in turn may lead to further improvements in future programs.

Special Method-Theory Relationships

Three areas of theory and methodology have special implications for each other: social environment and sample size, social norms and unit of assignment, and predictors of drug use and attrition.

Social Environment and Sample Size. From a methodological perspective, assigning only one or two schools (or other units such as communities) per experimental condition is of concern because of possible nonequivalence and confounds. Several smoking and drug use prevention studies suggest wide between-school variation in the rate of smoking and drug use by students. Although we have only limited understanding of the environmental causes of variation in substance use behavior, recent models of the onset process suggest that one's social environment is a very important determinant of drug use. Thus, both recent theory and data suggest that (1) more than one unit (school or classroom) should be assigned to each experimental condition of future studies; (2) samples should be sufficiently large to permit systematic exploration of classroom and school characteristics; and (3) more effort is

needed to understand the relationships between school or classroom characteristics and variation in drug use.

Social Norms and Unit of Assignment. Most school-based researchers assign whole schools to experimental conditions, but some assign classrooms within schools. Unit of assignment has other implications in addition to those addressed in the preceding paragraph. Psychosocial approaches to prevention might change the social norms of complete social systems; thus, in addition to improving relevant social skills, these programs might reduce the social pressures to use substances. No study has yet provided data to determine this. Studies that used within-school as well as between-school control groups could be informative. If programs do change norms for complete schools or classrooms, then multiple experimental conditions should not be assigned within schools. On the other hand, if program effects are mediated mostly by development of skills, then experimental conditions could be assigned within schools, classrooms, or any other social unit (although "contamination" of treatments could still be a problem). Tests are needed to determine whether psychosocial prevention programs of various types change norms in complete social systems (Best et al. 1984).

Predictors of Drug Use and Attrition. Attrition has been noted as a serious problem in longitudinal studies of school populations. Some work has been done on methods for testing whether the attrition experiences in any particular study relate to internal and external validity (Biglan and Ary 1985; Hansen et al. 1985) and on minimizing attrition (Pirie et al. 1989). However, the possible relationship between predictors of smoking onset and attrition has not yet received attention. For example, there is evidence that a major predictor of adolescent substance use is "rebelliousness." Although there has been little explication of the psychological processes involved, it seems clear that rebelliousness probably also predicts absenteeism and school dropout rate. This means that students at high risk of becoming drug users are the same students who are most likely to drop out. Future research needs to include assessments of the predictors of attrition; then if high-risk students are indeed dropping out of studies, further work will be needed to minimize such attrition or to find analytical approaches to adjust for it (Heckman 1976).

CONCLUSIONS

We have attempted to accomplish three things. First, we reviewed and integrated major predictors of drug use in five theoretical domains. Second, we established 12 ways in which theory is important to the enterprises of prevention program development, program evaluation, and science. Third, we derived the methodological implications of theory.

Theory is important. Without it, we would be even more lost than we have been to date in developing effective prevention programs; however, more diligent attempts to use theory will lead to more effective programs in the future. Without theory, our evaluations would have been even less useful than they have been; however, closer attention to theory will lead to more useful research and evaluations. Without theory, the science of prevention would not have advanced at all. Prevention theory has advanced significantly, particularly during the past decade, but it will advance more rapidly with greater use and application of existing theory and with further attempts to clarify, test, and improve existing theory.

NOTES

1. Jim Dwyer brought this view to our attention.
2. Few theorists address the issue of differential prediction of drug use onset and continuing use leading to abuse. The pattern of predictors for use of different substances or for use at different levels (experimentation/onset, continuing use/abuse) may differ (the specific factor model). For example, the relative strength of parent and peer influences probably varies as a function of both regular adolescent development and stage of adoption of substance use (Flay et al. 1983). In contrast, a common factor model suggests that the same factors predict lower or higher levels or involvement for all substances. These two views make different predictions and suggest different analyses (Hays et al. 1987). However, a common factor etiology model is not incompatible with a simplex model for stages of adoption of drug use. Each substance may have a separate threshold value or level (similar to item characteristic curves in psychological test theory) on an underlying substance use dimension. A low subject score on this dimension implies no substance use; a very high score implies use of all substances; and an intermediate score implies use of all substances whose thresholds are less than that score. Environmental and intrapersonal factors might predict this underlying score.
3. Other variables such as normative expectations, knowledge of social influences, and resistance skills did not change as hypothesized, perhaps due to poor measurement.
4. See Flay (1987) for a discussion of different approaches for assessing construct validity.

REFERENCES

Ajzen, I., and Fishbein, M. *Understanding Attitudes and Predicting Social Behavior.* Englewood Cliffs, NJ: Prentice-Hall, 1980.

Akers, R.L. *Deviant Behavior: A Social Learning Perspective.* Belmont, CA: Wadsworth, 1977.

Akers, R.L. A social behaviorist's perspective on integration of theories of crime and deviance. In: Messuer, S.F.; Krohn, M.D.; Liska, A.E., eds. *Theoretical Integration in the Study of Crime and Delinquency.* Albany, NY: State University of New York Press, 1989. pp. 102-104.

Akers, R.L., and Cochran, J.K. Adolescent marijuana use: A test of three theories of deviant behavior. *Deviant Behav* 6:323-346, 1983.

Akers, R.L.; Krohn, M.D.; Lanza-Kaduce, L.; and Radosevich, M.J. Social learning and deviant behavior: A specific test of a general theory. *Am Sociol Rev* 44:636-655, 1979.

Bandura, A. *Social Learning Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1977.

Bandura, A. *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1986.

Best, J.A.; Flay, B.R.; Towson, S.M.J.; Ryan, K.B.; Perry, C.L.; Brown, K.S.; Kersell, M.W.; and D'Avernas, J.R. Smoking prevention and the concept of risk. *J Appl Soc Psychol* 14(3):257-273, 1984.

Bickman, L. The functions of program theory. *New Dir Program Eval* 33:5-18, 1987.

Biglan, A., and Ary, D.V. Methodological issues in research on smoking prevention. In: Bell, C.S., and Battjes, R.J., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 170-195.

Biglan, A., and Lichtenstein, E. A behavior-analytic approach to smoking acquisition: Some recent findings. *J Appl Soc Psychol* 13:207-223, 1984.

Brannon, B.R.; Dent, C.W.; Flay, B.R.; Smith, G.; Sussman, S.; Hansen, W.B.; and Johnson, C.A. The Television, School, and Family Project: V. An examination of program acceptance by treatment modality. *Prev Med,* in press.

Braucht, G.N.; Brakarsh, D.; Follinstad, D.; and Berry, K.L. Deviant drug use in adolescence: A review of psychological correlates. *Psychol Bull* 79(2):92-106, 1973.

Brook, J.S.; Whiteman, M.; Gordon, A.S.; and Brook, D.W. Paternal determinant of female adolescent's marijuana use. *Dev Psychol* 20:1032-1043, 1984.

Brook, J.S.; Whiteman, M.; Gordon, A.S.; and Cohen, P. Dynamics of childhood and adolescent personality traits and adolescent drug use. *Dev Psychol* 22(3):403-414, 1985.

Campbell, D.T. Definitional versus multiple operationalism. *Et Al,* 2(1):4-17, 1969.

Chen, H.T., and Rossi, P.H. The multigoal, theory-driven approach to evaluation: A model linking basic and applied science. *Soc Forces* 59:106-122, 1980.

Chen, H.T., and Rossi, P.H. Evaluating with sense: The theory-driven approach. *Eval Rev* 7:283-302, 1983.

Cook, T.D., and Campbell, D.T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally, 1979.

Dwyer, J.H. *Statistical Models for the Social and Behavioral Sciences.* New York: Oxford University Press, 1983.

Elliott, D.S.; Huizinga, D.; and Ageton, S.S. *Explaining Delinquency and Drug Use.* Beverly Hills: Sage Publications, 1985.

Evans, R.I. Health promotion—science or ideology? *Health Psychol* 7:203-219, 1988.

Fishbein, M., and Ajzen, I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Reading, MA: Addison-Wesley, 1975.

Flay, B.R. On improving the chances of mass media health promotion programs causing meaningful changes in behavior. In: Meyer, M., ed. *Health Education by Television and Radio.* Munich: Saur, 1981. pp. 56-89.

Flay, B.R. Psychosocial approaches to smoking prevention: A review of findings. *Health Psychol* 4(5):449-488, 1985.

Flay, B.R. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 15:451-474, 1986.

Flay, B.R.; D'Avernas, J.R.; Best, J.A.; Kersell, M.W.; and Ryan, K.B. Cigarette smoking: Why young people do it and ways of preventing it. In: McGrath, P.J., and Firestone, P., eds. *Pediatric and Adolescent Behavioral Medicine.* New York: Springer-Verlag, 1983. pp. 132-183.

Glasgow, R.E., and McCaul, K.D. Social and personal skills training programs for smoking prevention: Critique and directions for future research. In: Bell, C.S., and Battjes, R.J., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 50-66.

Goodstadt, M.S. Drug education—A turn-on or a turn-off? *J Drug Educ* 10:89-98, 1980.

Gorsuch, R.L., and Butler, M.C. Initial drug abuse: A review of predisposing social psychological factors. *Psychol Bull* 83(1):120-137, 1976.

Hansen, W.B.; Collins, L.M.; Malotte, C.K.; Johnson, C.A.; and Fielding, J.E. Attrition in prevention research. *J Behav Med* 8(3):261-275, 1985.

Hawkins, J.D.; Lishner, D.M.; and Catalano, R.F., Jr. Childhood predictors and the prevention of adolescent substance abuse. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 75-126.

Hawkins, J.D.; Lishner, D.M.; Catalano, R.F.; and Howard, M.O. Childhood predictors of adolescent substance abuse: Toward an empirically grounded theory. *J Child Contemp Soc* 8:11-48, 1986.

Hawkins, J.D., and Weis, J.G. The social development model: An integrated approach to delinquency prevention. *J Primary Prev* 6:73-97, 1985.

Hays, R.D.; Widaman, K.F.; DiMatteo, M.R.; and Stacy, A.W. Structural equation models of current drug use: Are appropriate models so simple (x)? *J Pers Soc Psychol* 52:134-144, 1987.

Heckman, J. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. *Ann Econ Soc Meas* 5:475-492, 1976.

Homans, G.C. *The Nature of Social Science.* New York: Harcourt, Brace & World, 1967.

Huba, G.J., and Bentler, P.M. The role of peer and adult models for drug taking at different stages in adolescence. *J Youth Adolescence* 9:449-465, 1980.

Huba, G.J., and Bentler, P.M. A developmental theory of drug use: Derivation and assessment of a causal modeling approach. In: Baltes, B.P., and Brim, O.G., Jr., eds. *Life-Span Development and Behavior.* Vol. 4. New York: Academic Press, 1982.

Huba, G.J.; Wingard, J.A.; and Bentler, P.M. Framework for an interactive theory of drug use. In: Lettieri, D.J.; Sayers, M.; and Pearson, H.W., eds. *Theories on Drug Abuse: Selected Contemporary Perspectives.* National Institute on Drug Abuse Research Monograph 30. DHHS Pub. No. (ADM)84-967. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1984. pp. 95-101.

Jessor, R. Marijuana: A review of recent psychosocial research. In: Dupont, R.L.; Goldstein, A.; and O'Donnell, J., eds. *Handbook on Drug Abuse.* Washington, DC: National Institute on Drug Abuse, 1979.

Jessor, R.; and Jessor, S.L. *Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth.* New York: Academic Press, 1977.

Johnston, L.D.; Bachman, J.G.; and O'Malley, P.M. *Student Drug Use, Attitudes, and Beliefs: National Trends, 1975-1982.* Rockville, MD: National Institute on Drug Abuse, 1982.

Judd, C.G., and Kenney, D.A. Process analysis: Estimating mediation in treatment evaluations. *Eval Rev* 5:602-619, 1981.

Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere, 1978a.

Kandel, D.B. Convergences in prospective longitudinal surveys of drug use in normal populations. In: Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere, 1978b.

Kandel, D.B. Drug and drinking behavior among youth. *Annu Rev Sociol* 6:235-285, 1980.

Kandel, D.B. Epidemiological and psychosocial perspectives on adolescent drug use. *J Am Acad Clin Psychiatry* 21:328-347, 1982.

Kandel, D.B.; Kessler, R.C.; and Margulies, R.Z. Antecedents of adolescents initiation into stages of drug use: A developmental analysis. In: Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere, 1978.

Kaplan, H.B. Increase in self-rejection as an antecedent of deviant responses. *J Youth Adolescence* 4(3):281-292, 1975.

Kaplan, H.B.; Martin, S.S.; and Robbins, C. Application of a general theory of deviant behavior: Self-derogation and adolescent drug use. *J Health Soc Behav* 23:274-294, 1982.

Kaplan, H.B.; Martin, S.S., and Robbins, C. Pathways to adolescent drug use: Self-derogation, peer influence, weakening of social controls, and early substance use. *J Health Soc Behav* 25:270-289, 1984.

Kellam, S.G.; Brown, C.H.; and Fleming, J.P. The prevention of teenage substance use: Longitudinal research and strategy. In: Coates, T.J.; Peterson, A.C.; and Perry, C., eds. *Promoting Adolescent Health: A Dialog on Research and Practice.* New York: Academic Press, 1982. pp. 171-200.

Lettieri, D.J., and Ludford, J.P., eds. *Drug Abuse and the American Adolescent.* National Institute on Drug Abuse Research Monograph 38. DHHS Pub. No. (ADM)81-1166. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1981.

Lettieri, D.J.; Sayers, M.; and Pearson, H.W., eds. *Theories on Drug Abuse: Selected Contemporary Perspectives.* National Institute on Drug Abuse Research Monograph 30. DHHS Pub. No. (ADM)84-967. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1984.

Lipsey, M.W.; Crosse, S.; Dunkle, J.; Pollard, J.; and Stobart, G. Evaluation: The state of the art and the sorry state of the science. *New Dir Program Eval* 27:7-28, 1985.

MacKinnon, D.P.; Johnson, C.A.; Pentz, M.A.P.; Dwyer, J.H.; Hansen, W.B.; Flay, B.R.; and Wang, E. "School-Based Prevention Effects on Mediating Variables: One-Year Results of the Midwestern Prevention Project." Unpublished manuscript, University of Southern California.

McCaul, K.D., and Glasgow, R.E. Preventing adolescent smoking: What have we learned about treatment construct validity? *Health Psychol* 4:361-387, 1985.

Merton, R.K. *Social Theory and Social Structure.* Glencoe, IL: Free Press, 1957.

Miller, J.D.; Cisin, I.H.; Gardner-Keaton, H.; Harrell, A.V.; Wirtz, P.W.; Abelson, H.I.; and Fishburne, P.M. *National Survey on Drug Abuse: Main Findings 1982.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)83-1268. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983.

Murray, D.M., and Perry, C.L. The prevention of adolescent drug abuse: Implications of etiological, developmental, behavioral, and environmental models. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 236-256.

Newcomb, M.D., and Bentler, P.M. *Consequences of Adolescent Drug Use: Impact on the Lives of Young Adults.* Beverly Hills: Sage Publications, 1988.

Newcomb, M.D.; Huba, G.J.; and Bentler, P.M. Mothers' influence on the drug use of their children: Confirmatory tests of direct modeling and mediational theories. *Dev Psychol* 19:714-726, 1983.

Pentz, M.A.; Brannon, B.R.; Charlin, V.L.; Barrett, E.J.; MacKinnon, D.P.; and Flay, B.R. The power of policy: Relationship of smoking policy to adolescent smoking. *Am J Public Health,* in press.

Pirie, P.L.; Murray, D.M.; Peterson, A.V.; Thompson, S.J.; Mann, S.L.; and Flay, B.R. Tracking and attrition in longitudinal school-based prevention research. *Prev Med* 18:249-256, 1989.

Richardson, J.; Dwyer, K.; McGuigan, K.; Hansen, W.B.; Dent, C.; Johnson, C.A.; Sussman, S.Y.; Brannon, B.; and Flay, B.R. Substance use among eighth graders who take care of themselves after school. *Pediatrics* 84:556-566, 1989.

Rudner, R.S. *Philosophy of Social Science.* Englewood Cliffs, NJ: Prentice-Hall, 1966.

Sadava, S.W. Psychosocial interactions and substance use. *Drugs Soc* 2(1):1-30, 1987.

Sadava, S.W., and Forsyth, R. Person-environment interaction and college student drug use: A multivariate longitudinal study. *Genet Psychol Monogr* 96:211-245, 1977.

Sechrest, L.; West, S.B.; Phillips, M.A.; Redner, R.; and Yeaton, W. Some neglected problems in evaluation research: Strength and integrity of treatments. In: Sechrest, L.; West, S.B.; Phillips, M.A.; Redner, R.; and Yeaton, W., eds. *Evaluation Studies Review Annual.* Vol. 4. Beverly Hills: Sage Publications, 1979.

Shiffman, S., and Wills, T.A. *Coping and Substance Use.* Orlando, FL: Academic Press, 1985.

Shontz, F.C., and Spotts, J.V. Who are the drug users? *Drugs Soc* 1:51-74, 1986.

Shore, M.F. Correlates and concepts: Are we chasing our tails? In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 127-135.

Smith, G.M., and Fogg, C.P. Psychological predictors of early use, late use, and non-use of marijuana among teenage students. In: Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere, 1978.

Sobol, D.F.; Rohrbach, L.A.; Dent, C.W.; Gleason, L.; Brannon, B.; Johnson, C.A.; and Flay, B.R. The integrity of smoking prevention curriculum delivery. *Health Educ Res* 4(1):59-68, 1989.

Suchman, E.A. *Evaluative Research: Principles and Practice in Public Service and Social Action Programs.* New York: Russell Sage Foundation, 1967.

Sussman, S.; Dent, C.W.; Brannon, B.R.; Glowacz, K.; Gleason, L.R.; Hansen, W.B.; Johnson, C.A.; and Flay, B.R. The television, school, and family smoking prevention/cessation project: IV. Controlling for program success expectancies across experimental and control conditions. *Addict Behav* 14(6):601-610, 1989.

Weiss, C.H. *Evaluation Research: Methods for Assessing Program Effectiveness.* Englewood Cliffs, NJ: Prentice-Hall, 1972.

Wilson, W.J. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago: University of Chicago Press, 1987.

Wingard, J.A.; Huba, G.J.; and Bentler, P.M. The relationship of personality structure to patterns of adolescent substance use. *Multivariate Behav Res* 14:131-143, 1979.

Wingard, J.A.; Huba, G.J.; and Bentler, P.M. A longitudinal analysis of personality structure and adolescent substance use. *Pers Individ Diff* 1:259-272, 1980.

Winick, C. A sociological theory of the genesis of drug dependence. In: Winick, C., ed. *Sociological Aspects of Drug Dependence.* Cleveland, OH: CRC Press, 1974.

Winick, C. The deviance model of drug-taking behavior: A critique. *Drugs Soc* 1:29-50, 1986.

BIBLIOGRAPHY—TABLE 1

Bailey, S.L., and Hubbard, R.L. Developmental variation in the context of marijuana initiation among adolescents. *J Health Soc Behav* 31:58-70, 1990.

Baumrind, D. Familial antecedents of adolescent drug use: A developmental perspective. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off, 1985. pp. 13-44.

Block, J., and Block, J.H. Longitudinally foretelling drug usage in adolescence: Early childhood personality and environmental precursors. *Child Dev* 59:403-414, 1988.

Block, J., Block, J.H., and Keyes, S. Longitudinally foretelling drug usage in adolescence: Early childhood personality and environmental precursors. *Child Dev* 59:403-414, 1988.

Brook, J.S.; Whiteman, M.; Gordon, A.S.; and Cohen, P. Dynamics of childhood and adolescent personality traits and adolescent drug use. *Dev Psychol* 22(3):403-414, 1985.

Brook, J.S.; Whiteman, M.; Gordon, A.S.; Nomura, C.; and Brook, D.W. Onset of adolescent drinking: A longitudinal study of intrapersonal and interpersonal antecedents. *Adv Alcohol Subst Abuse* 5(3):91-110, 1986.

Campbell, E.Q. The internalization of moral norms. *Sociometry* 27(4):391-412, 1964.

Christiansen, B.A.; Smith, G.T.; Roehling, P.V.; and Goldman, M.S. Using alcohol expectancies to predict adolescent drinking behavior after one year. *J Consult Clin Psychol* 57(1):93-99, 1989.

Donovan, J.E., and Jessor, R. Structure of problem behavior in adolescence and young adulthood. *J Consult Clin Psychol* 52(6):890-904, 1985.

Donovan, J.E.; Jessor, R.; and Jessor, L. Problem drinking in adolescence and young adulthood: A followup study. *J Stud Alcohol* 44(1):109-137, 1983.

Elliott, D.S.; Huizinga, D.; and Ageton, S.S. *Explaining Delinquency and Drug Use.* Beverly Hills: Sage Publications, 1985.

Friedman, A.S.; Utada, A.T.; Glickman, N.W.; and Morrissey, M.R. Psychopathology as an antecedent to, and as a "consequence" of, substance use in adolescence. *J Drug Educ* 17:233-244, 1987.

Galambos, N.L., and Silbereisen, R.K. Influence of income change and parental acceptance on adolescent transgression proness and peer relations. *Eur J Psychol Educ* 1:17-28, 1987.

Huba, G.J., and Bentler, P.M. A developmental theory of drug use: Derivation and assessment of a causal modeling approach. In: Baltes, B.P., and Brim, O.G., Jr., eds. *Life-Span Development and Behavior.* Vol. 4. New York: Academic Press, 1982.

Huba, G.J., and Bentler, P.M. Causal models of the development of law abidance and its relationship to psychosocial factors and drug use. In: Laufer, W.S., and Day, J.M., eds. *Personality, Moral Development, and Criminal Behavior.* Lexington, MA: D.C. Heath and Company, 1983. pp. 165-215.

Huba, G.J., and Bentler, P.M. Causal models of personality, peer culture characteristics, drug use, and criminal behavior over a five-year span. In: Goodwin, D.W.; Van Dusen, K.T.; and Mednick, S.A., eds. *Longitudinal Research in Alcoholism.* Boston: Klower-Nijhof, 1984. pp. 73-94.

Huba, G.J.; Newcomb, M.D.; and Bentler, P.M. Adverse drug experiences and drug use behaviors: A one-year longitudinal study of adolescents. *J Pediatr Psychol* 11(2):203-219, 1986.

Huba, G.J.; Wingard, J.A.; and Bentler, P.M. Longitudinal analysis of the role of peer support, adult models, and peer subcultures in the beginning adolescent substance use: An application of setwise canonical correlation methods. *Multivariate Behav Res* 15:259-279, 1980.

Huba, G.J.; Wingard, J.A.; and Bentler, P.M. Intentions to use drugs among adolescents: A longitudinal analysis. *Int J Addict* 16:331-339, 1981.

Jessor, R. Predicting time of onset of marijuana use: A developmental study of high school youth. *J Consult Clin Psychol* 44(1):125-134, 1976.

Jessor, R., and Jessor, S.L. *Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth.* New York: Academic Press, 1977.

Jessor, R.; Jessor, S.L.; and Finney, J. A social psychology of marijuana use: Longitudinal studies of high school and college youth. *J Pers Soc Psychol* 26(1):1-15, 1973.

Johnston, L.D. *Drugs and American Youth.* Ann Arbor, MI: University of Michigan Press, 1973.

Johnston, L.D. Drug use during and after high school: Results of a national longitudinal study. *Am J Public Health* 64[Suppl]:29-37, 1974.

Johnston, L.D.; O'Malley, P.M.; and Eveland, L.K. Drugs and delinquency: A search for causal connections. In: Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere, 1978.

Kandel, D.B.; Kessler, R.C.; and Margulies, R.Z. Antecedents of adolescent initiation into stages of drug use: A developmental analysis. In: Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere, 1978.

Kandel, D.; Simcha-Fagan, O.; and Davies, M. Risk factors for delinquency and illicit drug use from adolescence to young adulthood. *J Drug Issues* 16(1):67-90, 1986.

Kaplan, H.B. Increase in self-rejection as an antecedent of deviant responses. *J Youth Adolescence* 4(3):281-292, 1975.

Kaplan, H.B.; Martin, S.S.; Johnson, R.J.; and Robbins, C.A. Escalation of marijuana use: Application of a general theory of deviant behavior. *J Health Soc Behav* 27:44-61, 1986.

Kaplan, H.B.; Martin, S.S.; and Robbins, C. Application of a general theory of deviant behavior: Self-derogation and adolescent drug use. *J Health Soc Behav* 23:274-294, 1982.

Kaplan, H.B.; Martin, S.S.; and Robbins, C. Pathways to adolescent drug use: Self-derogation, peer influence, weakening of social controls, and early substance use. *J Health Soc Behav* 25:270-289, 1984.

Kellam, S.G.; Brown, C.H.; and Fleming, J.P. The prevention of teenage substance use: Longitudinal research and strategy. In: Coates, T.J.; Peterson, A.C.; and Perry, C., eds. *Promoting Adolescent Health: A Dialog on Research and Practice.* New York: Academic Press, 1982. pp. 171-200.

Kellam, S.G.; Simon, M.B.; and Ensminger, M.E. Antecedents in first grade of teenage drug use and psychological well-being: A ten-year communitywide prospective study. In: *Prevention of Drug Abuse.* Hearings before the Select Committee on Narcotics Abuse and Control, House of Representatives, Ninety-Fifth Congress, Second Session, April 18, 20, 25; May 16, 25. Washington, DC: U.S. Govt. Print. Off., 1978 (SCNAC-95-2-4).

Maddahian, E.; Newcomb, M.D.; and Bentler, P.M. Single and multiple patterns of adolescent drug substance use: Longitudinal comparisons of four ethnic groups. *J Drug Educ* 15:311-326, 1985.

Maddahian, E.; Newcomb, M.D.; and Bentler, P.M. Adolescents' substance use: Impact of ethnicity, income, and availability. *Adv Alcohol Subst Abuse* 5(3):63-78, 1986.

Maddahian, E.; Newcomb, M.D.; and Bentler, P.M. Adolescent drug use and intention to use drugs: Concurrent and longitudinal analyses of four ethnic groups. *Addict Behav* 13:191-195, 1988.

Margulies, R.Z.; Kessler, R.C.; and Kandel, B.D. A longitudinal study of onset of drinking among high school students. *J Stud Alcohol* 38(5):897-912, 1977.

Newcomb, M.D. *Drug Use in the Workplace: Risk Factors for Disruptive Substance Use Among Young Adults.* Dover, MA: Auburn House Publishing Company, 1988.

Newcomb, M.D., and Bentler, P.M. Cocaine use among adolescents: Longitudinal associations with social context, psychopathology, and use of other substances. *Addict Behav* 11:263-273, 1986.

Newcomb, M.D., and Bentler, P.M. Drug use, educational aspirations, and work force involvement: The transition from adolescence to young adulthood. *Am J Community Psychol* 14:303-321, 1986.

Newcomb, M.D., and Bentler, P.M. Frequency and sequence of drug use: A longitudinal study from early adolescence to young adulthood. *J Drug Educ* 16:101-120, 1986.

Newcomb, M.D., and Bentler, P.M. Changes in drug use from high school to young adulthood: Effects of living arrangement and current life pursuit. *J Appl Dev Psychol* 8:221-246, 1987.

Newcomb, M.D., and Bentler, P.M. Impact of adolescent drug use and social support on problems of young adults: A longitudinal study. *J Abnorm Psychol* 97(1):64-75, 1988.

Newcomb, M.D., and Bentler, P.M. The impact of family context, deviant attitudes, and emotional distress on adolescent drug use: Longitudinal latent-variable analyses of mothers and their children. *J Res Pers* 22:154-176, 1988.

Newcomb, M.D., and Harlow, L.L. Life events and substance use among adolescents: Mediating effects of perceived loss of control and meaninglessness in life. *J Pers Soc Psychol* 51:564-577, 1986.

Newcomb, M.D.; Maddahian, E.; and Bentler, P.M. Risk factors for drug use among adolescents: Concurrent and longitudinal analyses. *Am J Public Health* 76(5):525-531, 1986.

Newcomb, M.D.; Maddahian, E.; Skager, R.; and Bentler, P.M. Substance abuse and psychosocial risk factors among teenagers: Associations with sex, age, ethnicity, and type of school. *Am J Alcohol Abuse* 13(4):413-433, 1987.

Paton, S.; Kessler, R.; and Kandel, D.B. Depressive mood and adolescent illicit drug use: A longitudinal analysis. *J Genet Psychol* 131:267-289, 1977.

Pulkkinen, L. Youthful smoking and drinking in a longitudinal perspective. *J Youth Adolescence* 12(4):253-283, 1983.

Schlegel, R.P.; Manske, S.R.; and D'Avernas, J.R. Alcohol and drug use in young adults: Selected findings in a longitudinal study. *Addict Behav* 4:213-225, 1985.

Shedler, J., and Block, J. Adolescent drug use and psychological health: A longitudinal inquiry. *Am Psychol* 45:612-630, 1990.

Smith, G.M., and Fogg, C.P. Psychological predictors of early use, late use, and non-use of marijuana among teenage students. In: Kandel, D.B., ed. *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* Washington, DC: Hemisphere, 1978.

Stein, J.A.; Newcomb, M.D.; and Bentler, P.M. The relationship of gender, social conformity, and substance use: A longitudinal study. *Bull Soc Psychol Addict Behav* 5:125-138, 1986.

Stein, J.A.; Newcomb, M.D.; and Bentler, P.M. An 8-year study of multiple influences on drug use and drug use consequences. *J Pers Soc Psychol* 53(6):1094-1105, 1987.

Stein, J.A.; Newcomb, M.D.; and Bentler, P.M. Personality and drug use: Reciprocal effects across four years. *Pers Individ Diff* 8(3):419-430, 1987.

Teichman, M.; Zipora, B.; and Giora, R. Personality and substance use among adolescents: A longitudinal study. *Br J Addict* 14:181-190, 1989.

Vicary, J.R., and Lerner, J.V. Longitudinal perspectives on drug use: Analyses from the New York longitudinal study. *J Drug Educ* 13(3):275-285, 1983.

Windel, M. A longitudinal study of antisocial behaviors in early adolescence as predictors of late adolescence substance use: Gender and ethnic group differences. *J Abnorm Psychol* 99:86-91, 1990.

Yamaguchi, K., and Kandel, D.B. On the resolution of role of incompatibility: A life event history analysis of family and roles and marijuana use. *Am J Sociol* 90:1284-1325, 1985.

ACKNOWLEDGMENT

This work was supported by National Institute on Drug Abuse grant DA-03468 and National Cancer Institute grant CA-44907.

AUTHORS

Brian R. Flay, Ph.D.
Associate Professor and Director

John Petraitis, M.A.
Assistant Professor

Prevention Research Center
University of Illinois at Chicago
School of Public Health
Suite 400
850 West Jackson Boulevard
Chicago, IL 60607

140141

Defining the Intervention and the Target Population

Lewayne D. Gilchrist

INTRODUCTION

A growing consensus among prevention researchers states that a single drug prevention strategy cannot be expected to reach or benefit all adolescents (Bell and Battjes 1985; Jones and Battjes 1985, p. 227). "Identifying effective prevention approaches also requires the ability to target programs—to identify which types of individuals are effectively reached with a specific approach" (Battjes and Bell 1985). Axiomatic in textbooks on clinical intervention is the idea that outcome goals, considered together with feasibility issues, drive the choice of intervention target and method. In the field of drug prevention, however, there is a surprisingly unexamined lack of clarity in the definition of prevention goals and targets (Jessor 1984; Murray and Perry 1985). This lack of clarity substantially reduces the amount of useful information that can be gleaned from prevention trials and obscures potential new directions that might be taken toward developing a more unified, comprehensive, and useful science of prevention.

This chapter reviews factors related to selection of appropriate outcome goals for drug prevention research and also discusses some new directions for targeting preventive interventions. Finally, the chapter outlines suggestions for defining and reporting prevention studies so that results from prevention research contribute toward a more unified science of drug prevention.

DEFINING GOALS FOR PREVENTIVE INTERVENTIONS

At the present time, there is no well-articulated consensus in our field regarding goals for drug prevention programs. Prevention researchers are interested in theory and in testing methods suggested by theory for inducing desired behavior change. The goal of this effort—drug prevention—is often presumed to be self-evident or to reside in a distant political rather than scientific arena. Prevention research to date shows that sweeping changes in drug use among youth are unlikely to ensue from any single preventive program or approach.

Yet, in the current context of the war on drugs, drug prevention programs are expected to accomplish such broad-scale change. Unexamined and fundamentally unrealistic expectations regarding the power and impact of preventive programs can lead eventually to disillusion with the viability of prevention itself (Griffin 1986). An important contribution that prevention researchers can make is to help prevention research consumers (politicians, social scientists, professional practitioners, and the general public) understand the range of possible prevention goals and the appropriateness of each goal in the light of empirical findings.

In a comprehensive review of prevention literature, Hawkins and colleagues (unpublished manuscript) summarize at least seven different current views on appropriate end goals for drug prevention interventions:

1. Prevention programs should address eliminating patterns of pathological drug use that cause impairment in school and family settings, in overall social development, and in interpersonal relationships. Impairment can be defined as anything from school failure to lack of friends to depression and delinquency.

2. Prevention programs should reduce transition from experimental to repetitive use of drugs. Repetitive use is that which occurs with some designated frequency over a specified period—regardless of whether this use is accompanied by overt problems in personal, social, educational, or economic functioning. The goal is based on the assumption that it is regular, patterned use that leads to psychological dependency and to physiological addiction.

3. Prevention programing must eliminate any use of drugs regardless of whether this use is experimental, repetitive, persistent, or accompanied by social or personal problems. This is the abstinence goal, often a rallying cry for political and community groups.

4. Prevention programs should delay early onset of drug use. The goal of delaying initiation of drug use rests on research by Robins and Przybeck (1985) and others showing that youth who begin drug use before age 15 are more than twice as likely to develop drug abuse problems compared with those who initiate use at later stages of development.

5. Prevention efforts should delay or reduce initiation of so-called "gateway" substances, namely tobacco, alcohol, and marijuana, so that fewer adolescents will complete a sequence of transitions into use of increasingly "harder" and more harmful substances. Assumptions about the sequence

or stages of entry into hard drug use rest on research by Kandel and colleagues, who found stable patterns of progression from cigarettes, to alcohol, to marijuana, and to harder drugs for boys but less clear patterns of progression for girls (Kandel 1984; Yamaguchi and Kandel 1984).

6. Prevention efforts are most fruitfully focused on controlling circumstances involving drug use that may lead to immediate personal risk to self or others, such as driving drunk or engaging in unprotected sexual intercourse while high. This goal corresponds to Jessor's (1984) suggestions for focus on insulating youth from the harmful or irreversible consequences of drug use.

7. Prevention efforts must begin early in children's lives and must ameliorate major precursors of drug use. Empirical evidence underscoring this goal is as yet indirect. Research in progress (Hawkins and Catalano 1988; Hawkins et al., unpublished manuscript) has found that interventions to increase inner-city elementary school children's bonding to school and other positive social influences decrease delinquent behavior in upper elementary grade children. Theory predicts that increases in attachment to these values and decreases in delinquent behavior will lead to decreases in drug abuse. The study to verify this prediction is still under way.

It is possible to think of additional goals for programs aimed at preventing harm from drug use; for example, prevention of needle use among drug-using (but not needle-using) groups and prevention of drug dealing as opposed to drug using. In response to the political press to "do something about drugs," prevention programs are often launched without clear specification of a precise and well-justified end goal, thus generating imprecise notions about what the program can be expected to accomplish. Future prevention research should critically examine the validity, feasibility, and implications of alternative prevention goals and carefully (i.e., empirically) justify *a priori* the choice—both explicit and implicit—of preventive goal (anticipated outcome) for every preventive intervention that is tested.

Current literature contains information helpful for evaluating and prioritizing competing preventive intervention goals. With regard to the goal of preventing cigarette, alcohol, and marijuana use *per se* because these substances are gateways to use of harder drugs, recent findings from programs sponsored by the Office for Substance Abuse Prevention show that these three substances are not always the first steps in a steady progression to use of harder and harder drugs (McColgan 1989). Physical availability, cost, and cultural and peer norms can affect drug of initiation. In some communities, crack or inhalants can be the first drug used. Patterns of drug use vary geographically and over time.

It seems unlikely that one (or even three) specific substance(s) would remain the most salient intervention target(s) in all settings and at all periods.

In one of the few studies available on the consequences of using drugs in adolescence, Newcomb and Bentler (1988) show that harm in the form of long-term health, social, and economic damage in young adulthood resulted only from adolescents' regular, frequent, and committed drug use over long periods. Research by these investigators presents a persuasive argument in favor of focusing a science of drug prevention on eliminating acquisition of frequent, committed drug use. This overall goal encompasses the subgoals of delaying early onset of drug use and reducing the transition from experimental to repetitive use of drugs. Finally, the insulation-from-harm goal may be worth examining in more detail. Data from the Monitoring the Future surveys show that, given the wide availability of drugs in most communities, a great many youth may experiment with a variety of substances—especially cigarettes and alcohol and, to a lesser extent, marijuana—but that the majority do not move to regular, frequent, and committed use of any drug (Johnston et al. 1988). For the majority of adolescents, the most logical and efficient preventive program goal may be that of insulation from permanent or irrevocable consequences of temporary experimentation. More epidemiological research on the consequences of adolescent drug use is needed to assist the choice of an appropriate goal for specific communities and populations.

In summary, this kind of goal analysis and selection is rooted in empirical findings and suggests the need for most communities to address several different preventive goals simultaneously and, thus, for prevention researchers and community-based prevention planners to select several different preventive interventions to accomplish the overarching general goal of prevention of drug-precipitated harm to individuals and to communities.

DEFINING RISK

In the drug prevention field, definitions of who is at risk for regular, frequent, and committed drug use are no more stable or definitive than are definitions of preventive program goals. The ability to predict which adolescents are likely to become regular or habitual drug users has been widely accepted as important to cost-effective focusing of prevention efforts. Yet, etiological research to date has not produced any broadly useful tools for accurately assessing risk. Bypassing etiological research altogether, many school-based preventive interventions have simply assumed that all children and adolescents are at risk for substance use and, therefore, all youth are appropriate targets for preventive intervention. In fact, consensus exists in the drug field that adolescents differ with regard to the probability that they will become involved with drugs. There is

also consensus that drug use is multidetermined and that no single factor reliably "causes" problem substance use. Several researchers have reviewed etiologic studies that identify factors associated with or that appear predictive of drug use in adolescence and adulthood (Hawkins et al. 1985; Murray and Perry 1985; Perry and Jessor 1985). The following list, drawn from several such reviews, summarizes the variables that current etiologic studies identify as placing children and adolescents at higher risk for problem drug use.

- Individual Behavioral Factors

  —Academic failure
  —Early antisocial behavior
  —Early drug experimentation
  —Early drug use
  —Lack of behavioral skill

- Individual Attitudinal Factors

  —Rebelliousness against authority
  —Low commitment to school
  —Favorable attitudes toward deviance
  —Favorable attitudes toward adult behavior (transition proneness)

- Individual Psychological Factors

  —Low self-esteem
  —Low self-efficacy
  —Sensation seeking

- Family Environment Factors

  —Family history of drug use and/or antisocial behavior
  —Family management problems (low parenting skills)
  —High parental tolerance for deviance
  —Family disorganization

- Community Environment Factors

  —Economic and social deprivation
  —Community disorganization
  —Community norms favorable to deviance
  —Availability of drugs
  —Friends/peers who use drugs

Although some commonalities in pathways to problem drug use have been identified, findings from etiologic studies to date remain fragmented. Drug use is recognized as a complex and multidetermined phenomenon. Yet, etiologic studies often identify discrete risk variables that are not linked logically with each other or are linked in relatively simplistic ways. Further, etiologic studies often enumerate factors not amenable to intervention (i.e., race). Finally, the sheer number and variety of identified risk factors make it difficult to formulate a comprehensive, unified theory of risk and develop a valid and reliable risk assessment procedure.

Recognizing the complexity of drug abuse, Newcomb and Bentler (1988) have summarized four domains of variables that influence and that are influenced by drug use: biological (physiological processes), intrapersonal (within the individual), interpersonal (social), and sociocultural (community systems). Most tested preventive interventions to date have focused on interpersonal domain variables. As Newcomb and Bentler correctly point out,

> Focusing simply on handling peer pressure, such as the "just say no" approaches, may placate concerned but naive parents, teachers, and funding sources, but is an incomplete approach to confronting the task of preventing drug abuse among this Nation's youth (Newcomb and Bentler 1988, p. 234).

Even setting aside the issue of multiple domains of drug abuse determinants, etiologic findings to date do not provide much useful direction related to a series of variables that researchers increasingly recognize as critical to the receipt and impact of preventive—indeed all—interventions. These variables are *developmental status* (Murray and Perry 1985), *gender* (Gilchrist et al. 1989; Newcomb and Bentler 1988), *racial and cultural minority status* (National Institute on Alcohol Abuse and Alcoholism 1985), and *level of intentionality or actual drug involvement when the program begins* (Leventhal et al. 1985). All four of these factors serve as filters through which program recipients screen all clinical and educational interventions. Thus, these factors have great potential influence over the reaction of recipients to the preventive intervention's medium and message.

The biggest gap in extant etiologic research is lack of attention to the risk-precipitating role of the environment, in the sense of both interpersonal context and community norms, in onset, and in escalation of substance abuse. Literature from the field of drug treatment has long recognized that environmental and situational or contextual variables may exacerbate or may reduce risk of drug use among individuals who exhibit intrapersonal (trait, psychological, or cognitive-behavioral) vulnerability (Leventhal et al. 1985;

Marlatt and Gordon 1985). Well-articulated models of risk for drug use are needed that will unify individual/intrapersonal and community/environmental perspectives. Figure 1 is a graphic representation of a plausible distribution of individuals with regard to their vulnerability to drug use. The figure takes into account both individual and environmental/contextual factors that contribute to the probability that an individual will become a problem drug user. The remainder of this discussion addresses issues that may be helpful in devising goal selection and intervention targeting procedures for future research that take into account both individual and environmental/ contextual contributors to overall risk of drug use.

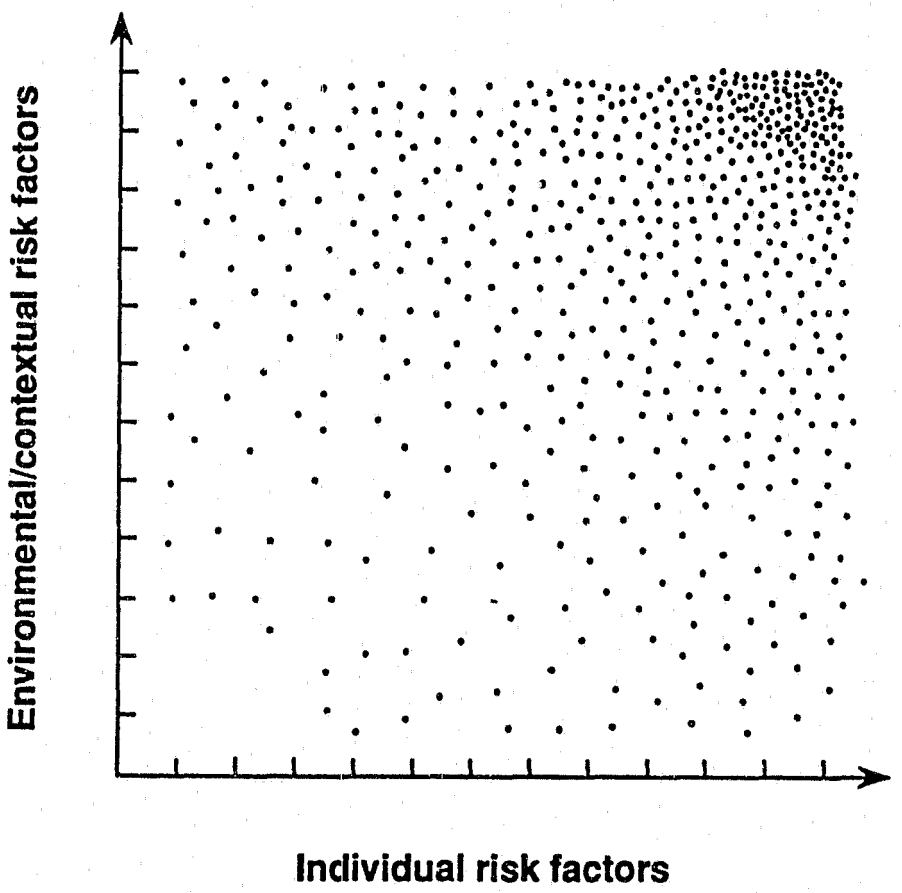

FIGURE 1. *Interaction of individual and environmental/contextual variables*

STEPS TOWARD A COMPREHENSIVE FRAMEWORK FOR SELECTING PREVENTION GOALS AND TARGETS

Several ideas drawn from recent literature may be helpful for defining preventive goals and intervention targets and enlarging both the scope and the precision of future drug prevention efforts.

Program Strength

If individuals differ in their vulnerability to drugs, it follows that they may require interventions of differing intensity to reach a given drug prevention-related goal. One framework for thinking about varied program intensity was outlined by Gordon (1983) and expanded more recently by others (Gilchrist 1990; Schinke et al. 1986). This framework is based on assessment of participants' risk potential and the probable cost-effectiveness of preventive program components. Preventive interventions may be said to be available in three "strengths"—universal, selective, and indicated.

Universal preventive strategies and interventions are those programs that can be advocated confidently for all adolescents. These prevention activities must be demonstrated to not be harmful to youth in any way, to benefit enough adolescents to be worthwhile, and to be acceptable and feasible under widely varying conditions. Universal prevention activities are thus broad in scope, relatively less strong or intensive, and less costly; but they also are the least likely among prevention strategies to yield enduring behavior change, particularly among those youth most at risk. Most school-based programs in skills-building and values clarification fall into this universal strategy category.

Selective prevention strategies are those interventions that for reasons of cost and potential burden to communities or participants can be recommended only when adolescents are members of a subgroup in which risk of becoming a drug user has been established as well above average, for example, sexual partners of drug abusers.

Finally, *indicated* prevention strategies are those interventions that are most intense and most costly to both participants and communities. As such, these strategies are reserved only for individually identified youth who are already members of a drug-using subgroup or who already manifest problem behaviors, including experimentation or regular use of one or more drugs.

Individual-In-Environment Assessment Techniques

If the model of multiple program strengths to be deployed to meet different preventive goals is to work, more and better methods are needed for assessing

both individual and environmental risk and, most important, for determining the interaction of individual and environmental risk factors to more accurately predict the probability of future problem drug use. Some community and neighborhood environments contain and produce more drug abusers than others. More effort should be made to develop valid and reliable criteria for determining the level of environmental risk (or drug promotiveness) in given communities and for determining which individuals within a community will be most and least resistant to predetermined levels of environmental risk. Once individual-in-environment risk status is determined (even crudely), selecting the needed strength and focus for preventive programing becomes clearer. Figure 2 illustrates a simple matrix for targeting efficient prevention programing goals, given four levels of predetermined individual-in-environment risk status.

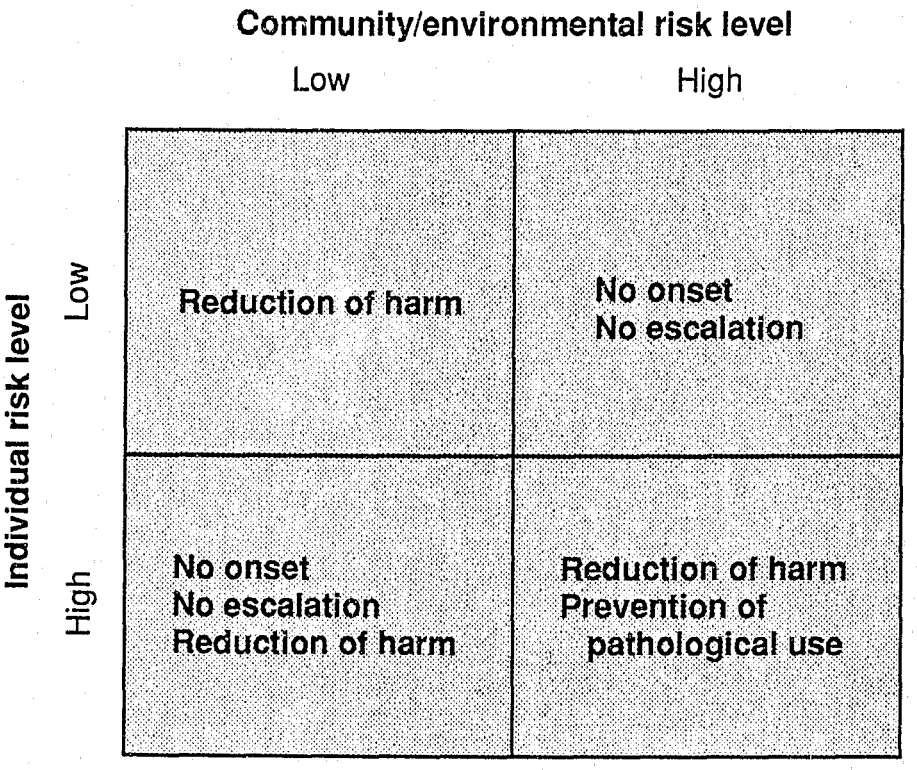

FIGURE 2. *Goal and target selection*

One conceptual task that is critical to achievement of improved assessment methods is empirical development of a definition of "community" that is workable for prevention programing purposes. At present, community can refer to anything from metropolitan area to city limits, school district, school, neighborhood, social or peer network, or individual family system. It is possible to think of each of these environmental layers (together with individuals) as potential units of intervention and develop testable hypotheses regarding interventions that focus simultaneously on two or more units.

Access Analysis for Program Delivery

Preventive interventions are delivered within the context of some system or community institution. Little has yet occurred in the drug prevention field in the way of systematic examination of opportunities beyond the school system for drug prevention interventions. Other systems, access, or program delivery opportunities exist, for example, health care systems, mental health care systems, law enforcement systems, and corporate or employment-related systems. Broadening beyond current school-based programing, table 1 depicts a matrix of system or access opportunities for focused preventive interventions that vary in intensity to meet varying levels of risk within a city or other large environmental unit.

TABLE 1. *System opportunities*

Program Strength	School	Health Care	Mental Health Care	Community Wide
UNIVERSAL All Sites Everybody	Information dissemination Resistance skills	Information dissemination Risk recognition and referral	Information dissemination Risk recognition and referral	Media campaigns for the general public Laws affecting norms
SELECTIVE Some Sites Everybody	Targeted programing for high-problem schools: To change school norms To involve families	Prevention programs in health clinics serving special populations: Pregnant adolescents Prospective parents Street youth	Prevention programs in community mental health care centers serving: Children of drug-using parents Children of depressed mothers	Detention programs Other youth programs Special media campaigns
INDICATED Some Sites Some Bodies	Individualized programs for special students	Treatment for drug-using parents to prevent drug-using children	Family therapy with referred families	Mandatory programs for drunk-driver arrestees

CONCLUSIONS

As a field, drug prevention has devoted insufficient attention to examining the range of possible preventive goals and their implications. In part, the field is not old enough to have reached this evolutionary stage. Many prevention research studies have been undertaken to shore up methodological holes in prior research and to reexamine existing behavior change technologies with improved research methods. Studies thus have built on one another while at the same time bypassing the issues of (1) finding clear and logical connections among etiologic findings, research on the consequences of drug use, environmental factors, institutional system or access opportunities, and cost-effectiveness and (2) welding these disparate factors into a comprehensive science of prevention that unites theory, individual and contextual risk assessment, access or program delivery opportunities, unit(s) of intervention, behavior change technologies, cost considerations, and specific outcome goals.

As the science of prevention develops, research can provide consumers with the information and assessment tools necessary to set realistic goals based on well-grounded expectations for what given prevention strategies and activities can and cannot do. At the very least, reports of all tests of preventive interventions should include precise definitions of the preventive goal, the empirically based rationale for selecting particular interventive targets (which involves the definition and justification of risk status), and a clear *a priori* prediction of the pattern of results that should be forthcoming from the intervention. Such studies need not be huge. The prevention field might fruitfully move from testing a single theoretical framework in large community settings to smaller scale, more focused examinations of interactions of program goal, program target, and program intensity with specified levels of risk found among adolescents accessed through defined community systems.

REFERENCES

Battjes, R.J., and Bell, C.S. Future directions in drug abuse prevention research. In: Bell, C.S., and Battjes, R.J., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 221-228.

Bell, C.S., and Battjes, R.J., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Gilchrist, L.D. The role of schools in community-based approaches to prevention of AIDS and intravenous drug use. In: Leukefeld, C.G.; Battjes,

R.J.; and Amsel, Z., eds. *AIDS and Intravenous Drug Use: Future Directions for Community-Based Prevention Research.* National Institute on Drug Abuse Research Monograph 93. DHHS Pub. No. 89-1627. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1990. pp. 150-166.

Gilchrist, L.D.; Schinke, S.P.; and Nurius, P. Reducing onset of habitual smoking among women. *Prev Med* 18:235-248, 1989.

Gordon, R.S., Jr. An operational classification of disease prevention. *Public Health Rep* 98:107-109, 1983.

Griffin, T. Community-based chemical use problem prevention. *J Sch Health* 56(9):414-417, 1986.

Hawkins, J.D., and Catalano, R.F. "Risk and Protective Factors for Alcohol and Other Drug Problems: Implications for Substance Abuse Prevention." Paper presented at the First Symposium on the Prevention of Alcohol and Other Drug Problems, Center of Alcohol Studies, Rutgers University, NJ, October 1988.

Hawkins, J.D.; Catalano, R.F.; Bridges, G.S.; Lake, L.; Gainey, R.; and Murphy, T. "A Risk-Based Analysis of Drug Abuse Prevention Strategies and Prospects." Unpublished manuscript.

Hawkins, J.D.; Lishner, D.M.; and Catalano, R.F., Jr. Childhood predictors and the prevention of adolescent substance abuse. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: U.S. Govt. Print. Off., 1985. pp. 75-126.

Jessor, R. Adolescent development and behavioral health. In: Matarazzo, J.D.; Weiss, S.M.; Herd, J.A.; Miller, N.E.; and Weiss, S.M., eds. *Behavioral Health: A Handbook of Health Enhancement and Disease Prevention.* New York: Wiley, 1984. pp. 69-90.

Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. *Illicit Drug Use, Smoking, and Drinking by America's High School Students, College Students, and Young Adults, 1975-1987.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)89-1602. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1988.

Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Kandel, D.B. Developmental stages in adolescent drug involvement. In: Lettieri, D.J.; Sayers, M.; and Pearson, H.W., eds. *Theories on Drug Abuse: Selected Contemporary Perspectives.* National Institute on Drug Abuse Research Monograph 30. DHHS Pub. No. (ADM)84-967. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1984. pp. 120-127.

Leventhal, H.; Prohaska, T.R.; and Hirschman, R.S. Preventive health behavior across the lifespan. In: Rosen, J.C., and Solomon, L.J., eds. *Prevention in*

Health Psychology. Hanover, NH: University Press of New England, 1985. pp. 191-235.

Marlatt, G.A., and Gordon, J.R. *Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors.* New York: Guilford, 1985.

McColgan B.R. Office for Substance Abuse Prevention. *OSAP High-Risk Youth Update,* April 1, 1989. p. 1.

Murray, D.M., and Perry, C.L. The prevention of adolescent drug abuse: Implications of etiological, developmental, behavioral, and environmental models. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 236-256.

National Institute on Alcohol Abuse and Alcoholism. *Alcohol Use Among U.S. Ethnic Minorities.* National Institute on Alcohol Abuse and Alcoholism Research Monograph 18. DHHS Pub. No. (ADM)89-1435. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Newcomb, M.D., and Bentler, P.M. *Consequences of Adolescent Drug Use: Impact on the Lives of Young Adults.* Newbury Park, CA: Sage, 1988.

Perry, C.L., and Jessor, R. The concept of health promotion and prevention of adolescent drug use. *Health Educ Q* 12(2):169-184, 1985.

Robins, L.N., and Przybeck, T.R. Age of onset of drug use as a factor in drug and other disorders. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 178-192.

Schinke, S.P.; Schilling, R.F. II; Gilchrist, L.D.; Whittaker, J.K.; Kirkham, M.A.; Senechal, V.A.; Snow, W.H.; and Maxwell, J.S. Definitions and methods for prevention research with youth and families. *Child Youth Serv Rev* 8:257-266, 1986.

Yamaguchi, K., and Kandel, D.B. Patterns of drug use from adolescence to young adulthood: II. Sequences of progression. *Am J Public Health* 74:668-672, 1984.

AUTHOR

Lewayne D. Gilchrist, Ph.D.
Associate Professor and Director
Center for Social Welfare Research
University of Washington School of Social Work
4101 15th Avenue, N.E.
Seattle, WA 98195

1H01H2

Implementation Issues in Drug Abuse Prevention Research

Mary Ann Pentz and Elizabeth Trebow

This chapter reviews methodological issues in evaluating the quality of implementation of drug abuse prevention programs. Issues of *definition* (adherence, exposure, reinvention), *measurement* (self-report, other's report, behavioral observation), and *parameters of influence* (person, situation, environment) are addressed. Implementation results of recent drug prevention and health promotion studies are reviewed as they relate to these issues. A general model is then proposed that represents implementation as a multiply determined process involving the interaction of person, situation, and environmental influences. Using this model, several recommendations are offered for estimating the "true" drug abuse prevention program effect as the average of effect estimates generated from experimental program assignment and level of program implementation. Potential differences between researcher and programer standards of quality or level of implementation are noted, using the common interpretation of the efficacy/effectiveness research trial continuum as an example.

INTRODUCTION

Reviews of drug abuse prevention research in the past decade indicate that primary prevention programs, particularly programs aimed at counteracting social influences to use drugs, have produced significant delays in the onset of smoking in young adolescents and lower rates of increase in the prevalence rates of gateway drugs—cigarettes, alcohol, and marijuana (Botvin 1986; Pentz 1983). Recently, however, prevention programs have been subject to criticism on several counts. First, effects have been variable, with some programs demonstrating effects on smoking but not on other drug use, some only short-term or only delayed effects, and others a decay in effect after a few years (Bangert-Drowns 1988; Battjes 1985). Second, the hypothesized mechanisms of behavior change in these programs do not often obtain, for example, the link between resistance skills taught in a social influences program and changes in drug use prevalence rates (Bangert-Drowns 1988; MacKinnon et al., submitted for publication). Third, there is little evidence to suggest that programs

developed from research studies are adopted by the consumers or are institutionalized by the organization in which the program was implemented after the researchers have left (Goodstadt 1988; Schaps et al. 1986). Fourth, the few research-generated programs that have been institutionalized tend to be modified substantially from the version used in research—for example, the DARE (Drug Abuse Resistance Education) program modified from Project SMART (Self-Management and Resistance Training) (Johnson 1986).

Collectively, these criticisms could be approached as questions about the quality of implementation of drug prevention programs. If systematically addressed in drug prevention research, analysis of program implementation might give a clearer answer to the question of whether drug prevention programs work. In addition, Basch (1984) has noted that implementation research, in general, can facilitate prevention program efforts by (1) elucidating factors that contribute to adoption and dissemination of programs, (2) expediting funding and resource allocation for programs, (3) enhancing the validity of summative evaluation derived from other measures, and (4) contributing to changes in theory and policy regarding programing. Unfortunately, much research on quality of implementation has been hampered by definitional problems, measurement problems, and a lack of consideration of predictors and mediators of implementation (Basch 1984; Fullan and Pomfret 1977; Leithwood and Montgomery 1980).

IMPLEMENTATION ISSUES

Definition

Program implementation can be conceptualized from at least three points of inquiry. Was the program implemented and/or implemented as designed *(adherence)*? How much of the program did consumers receive *(exposure)*? Was the program changed during implementation *(reinvention)*?

Adherence to experimental assignment is probably the definition of most interest to researchers (Basch et al. 1985). This definition can be specified in terms of whether the program and control groups adhered to their respective experimental conditions. Evidence of experimental crossover, contamination or unplanned diffusion of the program to control groups, and adoption of other conflicting programs by either the program or control groups would indicate lack of adherence (Cook and Campbell 1979; Cook 1985). Within program groups, adherence also can be measured as to whether the implementors report having used the program, whether the program was implemented with a strength sufficient to conclude that the program was delivered (e.g., if teaching 1 class of 10 is sufficient to conclude that the program was implemented), and whether

the program was implemented in a form or length such that consumers acknowledge receipt of the program. Whether the experimental assignment was carried out and whether the program was implemented at all are important questions for making valid conclusions about program effect. For example, if an efficacy trial showed that a prevention program was effective in slowing the rate of increase of drug use prevalence rates, but an effectiveness trial of the same program showed no effect, is the lack of effect in the latter trial due to poor research support and monitoring or was the program simply not implemented? The disadvantage of these basic interpretations of adherence is their lack of sensitivity to degrees of implementation or factors influencing implementation.

An alternative interpretation of adherence is program *fidelity*, or whether the program was implemented as it was originally designed by the researchers (Fullan and Pomfret 1977). This "by-the-book" interpretation requires that a "book" was developed, that is, that training and program materials and procedures are readily available as a research standard of the program. Fidelity measures how closely these materials were adhered to, either through subjective judgment of an evaluator who is familiar with the materials or through more objective documentation that specific procedures were completed. Typically, fidelity is operationalized as one or more of the following: (1) delivery of the requisite number of program sessions, activities, and tasks; (2) amount or frequency of on-task versus off-task behavior by the implementor; and/or (3) frequency of use of learning techniques used in training (e.g., modeling, rehearsal, discussion, and feedback vs. lecturing).[1] In implementation research, the question of fidelity is especially important for determining whether a program shown to be effective in an efficacy trial can show effects under effectiveness trial conditions, whether it is generalizable, and whether it could be replicated in other research studies. Fidelity is probably the most important concern of researchers who are interested in determining the maximum impact of a program. The potential disadvantage of focusing on a fidelity approach to implementation is that it assumes the "best" program effects will always derive from implementing the program exactly as it was designed by the researchers. Thus, fidelity ignores the contributions of reinvention and programer input to the magnitude and maintenance of program effects.

Reinvention is the extent to which program content and implementation are changed from the original standard as developed by the researchers and/or as agreed on by implementors and differs from lack of adherence in that it represents intentional or planned change (vs. lack of acceptance, noncooperation, or unplanned change) that is initiated for the purpose of enhancing program effectiveness. Measures of reinvention must differentiate planned changes in implementation from spontaneous historical events (e.g., school district changes in curriculum scheduling that affect the complete

implementation of a drug prevention program) and unplanned lack of adherence to the program (Cook 1985). Supplemental documentation often is required to validate that the change was planned (e.g., a formal review and rewrite of a prevention curriculum by researchers and program implementors). Diffusion of innovation theory suggests that, as a major indicator of adoption and institutionalization of a program (implementor "ownership"), reinvention should receive a high priority in evaluations of long-term effectiveness of drug abuse prevention programs (Rogers 1983). In addition to diffusion applications, reinvention is also important for determining whether program effects can be increased by tailoring content and implementation to certain environmental conditions or to certain populations of implementors and consumers. The tailoring question is particularly important to pursue in research with minority or high-risk populations when the original program is developed on white, middle-class populations.

Unfortunately, of all implementation constructs, reinvention is the most difficult to operationalize with a standardized measure (Basch 1984). Because the directions for reinvention may differ across program settings and populations, criterion-referenced tests may be more appropriate than standardized measures for evaluating whether reinvention occurred in drug prevention studies and for evaluating effects of reinvention on drug use behavior. In addition, more than other implementation constructs, measurement and analysis of reinvention also require assessing the parameters that influenced the decision to change a program. Because these parameters also may differ across program settings and populations, results of reinvention analyses may be difficult to generalize or replicate.

Measurement

In general, three types of measurement have been used to assess program implementation: *self-report, other's report,* and *behavioral observation.*

Self-report is the most commonly used implementation measure. In health promotion and disease prevention programs related to drug abuse prevention, most of which have been implemented in the school, implementor self-report (teacher or professional health educator) has been used extensively and often exclusively to measure implementation. Self-report measures have typically assessed program adherence and exposure. However, self-report measures have sometimes included assessment of program process (e.g., teacher perception of student disruptive behaviors), which represents a potential influence on program implementation and which may be used as a covariant in analyses of adherence and exposure. The major potential disadvantage of self-report is that it may be subject to response bias in the direction of social

desirability (Biglan and Ary 1985; Cook and Campbell 1979; Boruch and Gomez 1977). Thus, the quality of implementation based on self-report alone may be overestimated, and the potential magnitude of implementation effects on drug use behavior may be underestimated.

Other's report is assumed to be based directly on observation but also may be based indirectly on implementor self-report (e.g., teacher), consumer self-report (e.g., student), or interviews with individuals involved with either program implementation or planning (e.g., school administrator). The most commonly employed report is the research staff report, which typically focuses on adherence. However, several recent studies also have included student reports of program process and school administrator reports or archival records of teaching efficacy. Other's report typically correlates only moderately with self-report and is subject to the response bias of the reporter (Boruch and Gomez 1977).

Behavioral observation of program implementation by research staff or other observers who are independent of a program is considered the most objective measure of program implementation (Basch 1984). Observation usually is used to measure program fidelity, either as a sole source of this information or as an additional source to validate other measures of implementation. Until recently, behavioral observation was used in educational and behavior therapy research more than in prevention research. Several of the behavioral role-play situations and behavioral rating systems from these earlier studies have been adapted for use in drug prevention studies (Goldfried and Linehan 1977). For example, refusal and assertiveness skill role-plays have been adapted to drug use pressure situations, and ratings of behavioral states and events in classrooms have been adapted to assess the amount of time spent on drug abuse prevention role-plays vs. teacher lecturing about drug use. Consistent with findings in educational and behavior therapy research, behavioral observation shows only low-moderate relationships with self-reports and other's reports in drug prevention studies (Goldfried and Linehan 1977). Observation is also subject to different types of measurement problems (including observer effects on the implementor and consumer), possible reliance on ratings of behavior with low ecological validity (e.g., reliance on rating skills in a role-play that may never occur in real life), and difficulty and expense of use relative to self-report and other's report (Fullan and Pomfret 1977).

For the few studies that have compared different types of measures, little is known about why discrepancies in implementation ratings occur. Also, little is known about the relative relationships of each type of measure to drug use behavior outcomes.

Parameters of Influence or Predictors of Implementation

Most drug use epidemiological and prevention studies conducted in the past decade have evaluated predictors of drug use behavior. However, few of these studies have evaluated predictors of program implementation. Studies in education have routinely reported that contextual variables, including person, situation, and environmental variables and their interactions, predict student academic and social behavior (Burstein 1980; Fullan and Pomfret 1977; Raudenbush and Bryk 1986). These range from personal characteristics of the student that affect the acquisition of new skills, including achievement and learning ability test scores, to classroom environment represented by teacher/student and student/student interactions and to school and school district environment represented by types of courses offered and student enrollment. Also, behaviorally oriented therapy studies have indicated that therapeutic outcome (assumed to be directly related to quality of the therapy) is affected by the interaction of trainer (implementor), trainee, and therapy (program) characteristics (Pentz 1981). It appears logical that person, situation, and environment variables should affect the implementation of drug prevention programs as well (Perry and Murray 1985).

Person influences (intrapersonal-level variables) probably relate to the immediate quality of program implementation (e.g., session-by-session adherence and exposure) more than long-term outcomes associated with program reinvention and changes in drug use behavior. Using teacher-taught, school-based programs as a common example of drug prevention programs, person influences on implementation can be classified as teacher characteristics and student characteristics. Teacher characteristics include the ability to control students and the basic teaching skills. In social influences programs, active teaching skills of modeling, role-playing, and discussion demonstrated before training or program implementation are particularly important to assess. Current research suggests that these active teaching skills may be highly dependent on the comprehensiveness or length of training provided to the teacher and also may be negatively associated with teacher age. In terms of student characteristics, verbal and performance achievement levels have been associated with increases in assertiveness skill levels after program implementation, although the magnitude of change in skill levels may also depend on initial assertive, aggressive, or passive behaviors demonstrated by the student in social situations (Pentz 1981). Recent research also suggests that student expectancies may affect program implementation, although the direction of this relationship is still unclear (Sussman et al. 1989). Other research has indicated that comprehensiveness of peer leader training affects student behavior during implementation of social influences programs (Perry et al. 1988).

Situational influences (interpersonal) relate to immediate prevention skills transmission and acquisition, and they have been evaluated typically as program process. Situational influences on program implementation include the modeling, role-play, and discussion skills of implementors and students; the salience of the interpersonal interaction involving the demonstration of those skills (e.g., whether a drug use resistance situation being rehearsed by two students represents an actual situation they have encountered); and immediate competing interactions (e.g., student disruption of a role-play or high classroom noise levels during modeling).[2]

Environmental influences (extrapersonal level) probably relate to maintenance of program implementation over the long term and to program reinvention more than other influences. Environmental influences can be represented from the most proximal or microlevel that is expected to affect youth and program implementation to the more distal or macrolevels (Raudenbush and Bryk 1986). For example, in a school-based drug prevention program, the classroom environment might represent the most proximal environmental influence on program implementation and the school district environment the most distal. Classroom environment has been evaluated in previous studies; for example, those that have used the Moos' Classroom Environment Scale (CES) to measure student academic achievement, learning, and social behavior (Moos 1979). CES evaluates several factors of classroom environment that may relate directly to the quality of drug use prevention program implementation, including teacher/student and student/student relationships, class morale, and class learning motivation. In addition, implementation may be facilitated by implementor/student acceptance of the program (Brannon et al., in press). School environment can be measured by the frequency of disruption of school functioning (e.g., achievement testing, substitute teachers), teaching morale (e.g., teacher turnover), administrative support for teaching, and the number of courses or activities competing for the same schedule as a drug prevention program (Fors and Doster 1985; Goldstein 1984; Harnish 1987). At the school district/community level, influences on program implementation include district curriculum priorities and emphasis in the community on prevention vs. treatment and on demand vs. supply policies and interventions for drug use prevention (Moscowitz and Jones 1988; Newcomb et al. 1987; Pentz et al. 1986; Pentz et al. 1989).

RECENT PREVENTION PROGRAM IMPLEMENTATION STUDIES

Prevention implementation research studies are summarized in table 1. Most implementation research has concentrated on school-based health promotion and smoking or drug use prevention programs. The majority of studies have evaluated teachers as program implementors, used self-report measures of

TABLE 1. *Summary of prevention program implementation studies*

Study	Program	Implementor	Definition	Measure	Validation	Influences	Effects	Results
Life Skills Training (Botvin et al. 1983)	School drug prevention	Teacher	Exposure	Observation	No	—	Yes	Intensive exposure (massed vs. spaced sessions) and more sessions + related to smoking.
School Health Evaluation (SHEE) (Connell and Turner 1985; Connell et al. 1985; Fors and Doster 1985)	School health education	Teacher	Adherence (assignment) Fidelity Exposure Reinvention	Self-report	No	Person (teacher, prior instruction) Environment (school administrator support)	Yes	High adherence, fidelity, and exposure + related health knowledge, attitudes, skills, decreased smoking behavior. Teacher prior instruction + related to reinvention. Administrator support + related to fidelity and exposure.
Youth Health Promotion (Perry et al. 1988)	School drug prevention and smoking prevention	Student peer leaders	Fidelity (process)	Self-report Other's report	No	Person (student training) Environment (class support, family support)	No	Fidelity + related to training and peer support.
Alcohol Prevention Trial (Hansen et al., submitted for publication)	School alcohol prevention	Professional health educators	Fidelity	Self-report Observation	Yes	Environment (class enthusiasm)	No	Fidelity + related to alcohol prevention knowledge, perceived use norms, skills. Class enthusiasm, fidelity loaded on integrity index.

TABLE 1. *(Continued)*

Here's Looking at You, Two (Tricker and Davis 1988)	School drug prevention	Teacher	Adherence (assignment) Fidelity	Self-report	No	Person (teacher training) Environment (district financial support)	No	Adherence and fidelity + related to training and financial support.
Television, School, and Family Project (TVSFP) (Sobol et al. 1989)	School smoking prevention	Paraprofessional instructor	Fidelity	Self-report Videotape Observation	Yes	Situation (modeling, role-play, discussion skills)	No	Fidelity + related to person and situation-level skills.
Midwestern Prevention Project (MPP) (Pentz et al., submitted for publication)	School drug prevention	Teacher	Adherence (assignment) Fidelity Exposure Reinvention	Self-report Other's report Observation	Yes	Environment (school, SES, race, grade)	Yes	Adherence, fidelity, and exposure + related to decreased smoking, alcohol, and marijuana use.

implementation (with some validation by behavioral observation), and defined implementation as fidelity (a recent few have included exposure) to program content and teaching procedures. A few have evaluated the relative contribution of training to implementation and/or outcome. Only two studies have evaluated the direct relationship of implementation to drug use behavior. A few recent studies have adjusted for person influences in implementation analyses, including student demographic characteristics and expectancies and teacher experience, or situational influences represented by modeling or role-playing skills, but none has evaluated the predictive relationship of person, situational, and environmental influences to program implementation.

RECOMMENDATIONS

This chapter focuses on implementation issues, with specific applications to school-based drug use prevention programs, that are generalizable to other prevention programs as well. For example, for a grassroots, parent-based prevention program, the successive environmental influences would still progress from the microlevel environment in which the program is implemented to the macrolevel of environment, such as, from a small parent discussion group (similar to a classroom environment) to the parent-teacher association or other regional parent organizations to the community. The issues of definition and measurement are also generalizable across different types and sites for prevention programs. What *may* differ across programs are the specific implementation variables of interest or the components of the proposed model that are used to demonstrate program "success" (figure 1). For example, a drug use prevention program that is evaluated as part of an experimental research project should probably include all aspects of the model in assessment and analysis of program implementation. A program evaluation study, on the other hand, would probably place less emphasis on person influences (which may be fixed under "real-life" conditions), exposure, and types of measures used than on situational and environmental influences and adherence and reinvention. Variables selected from the model also will depend on the scope or funding of the study. A study that is funded primarily for evaluation of program effects on drug use behavior, for example, may not be structured to fund personnel for extensive program implementation observations. This concern is especially problematic in very large prevention studies. The design and funding limitations of most studies suggest that results of program implementation evaluation should be qualified in terms of the definition of implementation adopted for the study, the type(s) of measure employed, and the parameters of influence controlled for in analyses.

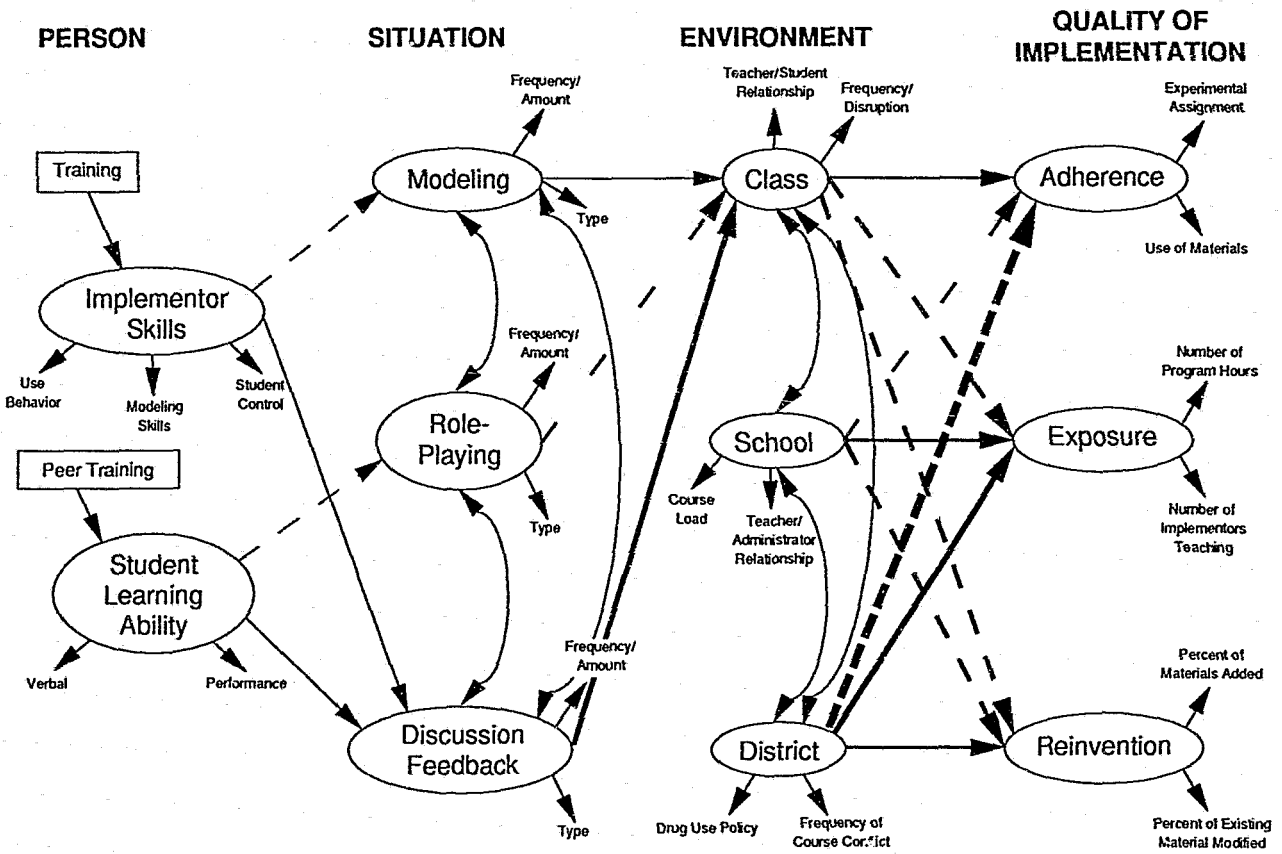

FIGURE 1. *Influences on quality of implementation, using a school prevention program as an example*

Given the increasing exigencies facing drug use prevention research at a time when national concern over the youth drug use problem is especially high, why should researchers attend to the specific issues of program implementation rather than concentrating their efforts on straightforward determination of program effects? The primary reason is that drug use prevention program effects may be significantly underestimated if experimental program assignment is used as the sole criterion or indicator of programing. Ironically, and somewhat contrary to the intended definition, using assignment as the criterion for program effect may represent an extreme case of effectiveness trial conditions, especially if a drug use prevention program is conducted under less supportive—although perhaps not less monitored—conditions than usual because the program has been controlled or "owned" by the researchers rather than by the school or community (Flay 1986; Pentz et al. 1986).

Future research and program evaluation studies might consider the possibility of two continua of efficacy-to-effectiveness trial conditions: the research/evaluation continuum and the program implementation continuum (figure 2). The first continuum provides a guideline for documenting whether the technology of the prevention program structure and process was followed as designed (Flay 1986), and the second provides a guideline for documenting the conditions under which the program structure and process are likely to have the maximum degree of support under real-life implementation conditions (Pentz et al. 1986). Considered from this perspective, program implementation may be a more logical and sensitive indicator of the actual effectiveness of a program.

Results of research on drug use prevention program implementation and outcome suggest that reliance on group assignment without regard to the quality of program implementation or the influences on program implementation probably yields a gross underestimate of program effectiveness. A more realistic estimate of "true" drug use prevention program effects may be generated from using the average of effect estimates generated from implementation and group assignment or, alternatively, developing confidence limits of program effectiveness bounded at the low end by program assignment and at the high end by implementation (Mark 1983).

NOTES

1. Adherence to the program as designed (program fidelity) is similar in concept to the "bandwidth of fidelity" of subjective measures, an area of research that has been recognized in the field of personality assessment research since the 1950s.

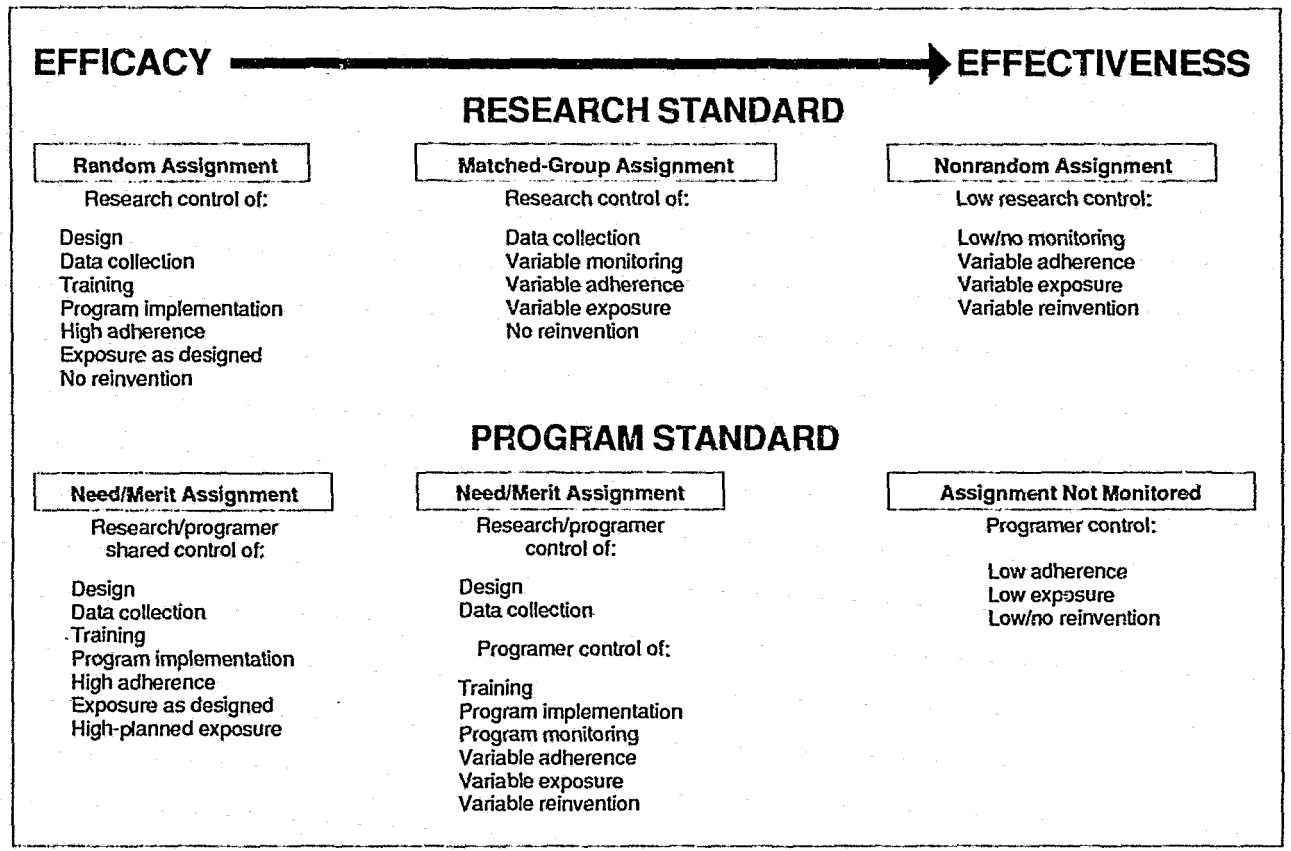

FIGURE 2. *Research and program standards for efficacy and effectiveness trials*

2. It could be argued that modeling, role-play, discussion, and other behavioral skills involving interpersonal interactions are indicators of implementation rather than predictors of implementation. However, the view presented in this chapter is that teaching/learning skills precede and determine adherence, exposure, and reinvention. For example, the basic modeling skills of a teacher will determine how much the teacher uses modeling of drug use resistance (if the latter is an implementation variable).

REFERENCES

Bangert-Drowns, R.L. The effects of school-based substance abuse education: A meta-analysis. *J Drug Educ* 18:246-264, 1988.

Basch, C.E. Research on disseminating and implementing health education programs in schools. *J Sch Health* 54:57-66, 1984.

Basch, C.E.; Sliepcevich, E.M.; Gold, R.S.; Duncan, D.F.; and Kolbe, L.J. Avoiding type III errors in health education program evaluations: A case study. *Health Educ Q* 12:315-331, 1985.

Battjes, R.J. Prevention of adolescent drug abuse. *Int J Addict* 20:1113-1134, 1985.

Biglan, A., and Ary, D.V. Methodological issues in research on smoking prevention. In: Bell, C.S., and Battjes, R.J., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 170-195.

Boruch, R.F., and Gomez, H. Sensitivity, bias, and theory in impact evaluations. *Professional Psychol* 8:411-434, 1977.

Botvin, G.J. Substance abuse prevention research: Recent developments and future directions. *J Sch Health* 56:369-374, 1986.

Botvin, G.J.; Renick, N.L.; and Baker, E. The effects of scheduling format and booster sessions on a broad-spectrum psychosocial approach to smoking prevention. *J Behav Med* 6:359-379, 1983.

Brannon, B.R.; Dent, C.W.; Flay, B.R.; Smith, G.; Sussman, S.; Johnson, C.A.; and Hansen, W.B. The Television, School, and Family Project: V. The impact of curriculum delivery format on program acceptance. *Prev Med*, in press.

Burstein, L. The role of levels of analysis in the specification of education effects. In: Dreeben, R., and Thomas, J.A., eds. *The Analysis of Educational Productivity. Vol. 1. Issues in Microanalysis.* Cambridge, MA: Ballinger Publishing Company, 1980. pp. 119-190.

Connell, D.B., and Turner, R.R. The impact of instructional experience and the effects of cumulative instruction. *J Sch Health* 55:324-331, 1985.

Connell, D.B.; Turner, R.R.; and Mason, E.F. Summary of findings of the school health education evaluation: Health promotion effectiveness, implementation, and costs. *J Sch Health* 55:316-321, 1985.

Cook, T.D. Priorities in research in smoking prevention. In: Bell, C.S., and Battjes, R.J., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 196-220.

Cook, T.D., and Campbell, D.T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally, 1979.

Flay, B.R. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 15:451-474, 1986.

Fors, S.W., and Doster, M.E. Implication of results: Factors for success. *J Sch Health* 55:332-334, 1985.

Fullan, M., and Pomfret, A. Research on curriculum and instruction implementation. *Rev Educ Res* 52:7-30, 1977.

Goldfried, M.R., and Linehan, M.M. Basic issues in behavioral assessment. In: Ciminero, A.R.; Calhoun, K.S.; and Adams, H.E., eds. *Handbook of Behavioral Assessment.* New York: Wiley Interscience, 1977. pp. 15-46.

Goldstein, H. The methodology of school comparisons. *Oxf Rev Educ* 10:69-74, 1984.

Goodstadt, M.J. School-based drug education in North America: What is wrong? What can be done? *J Sch Health* 58:278-281, 1988.

Hansen, W.B.; Graham, J.W.; Wolkenstein, B.H.; and Rohrbach, L.A. Program integrity as a moderator of prevention program effectiveness: Results for fifth grade students in the adolescent alcohol prevention trial, submitted for publication.

Harnish, D.L. Characteristics associated with effective public high schools. *J Educ Res* 80:233-241, 1987.

Johnson, C.A. Prevention and control of drug abuse. In: Last, J.M., ed. *Maxcy-Rosenau Public Health and Preventive Medicine.* Norwalk, CT: Appleton-Century-Crofts, 1986. pp. 1075-1087.

Leithwood, K.A., and Montgomery, D.J. Evaluating program implementation. *Eval Rev* 4:193-214, 1980.

MacKinnon, D.P.; Johnson, C.A.; Pentz, M.A.; Dwyer, J.H.; Hansen, W.B.; Flay, B.R.; and Wang, E.Y.I. Mediating mechanisms in a school-based drug prevention program: First year effects of the Midwestern Prevention Project, submitted for publication.

Mark, M.M. Treatment implementation, statistical power, and internal validity. *Eval Rev* 7:543-549, 1983.

Moos, R. *Evaluating Educational Environments: Procedures, Methods, Findings, and Policy Implications.* San Francisco: Jossey-Bass, 1979.

Moscowitz, J.M., and Jones, R. Alcohol and drug problems in the schools: Results of a national survey of school administrators. *J Stud Alcohol* 49:299-305, 1988.

Newcomb, M.D.; Maddahian, E.; Skager, R.; and Bentler, P.M. Substance abuse and psychosocial risk factors among teenagers: Associations with sex, age, ethnicity, and type of school. *Am J Drug Alcohol Abuse* 13:413-433, 1987.

Pentz, M.A. The contribution of individual differences to assertion training outcome in adolescents. *J Counseling Psychol* 28:529-532, 1981.

Pentz, M.A. Prevention of adolescent substance abuse through social skills development. *Preventing Adolescent Drug Abuse: Intervention Strategies.* National Institute on Drug Abuse Research Monograph 47. DHHS Pub. No. (ADM)83-1280. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983. pp. 195-231.

Pentz, M.A.; Cormack, C.; Flay, B.R.; Hansen, W.B.; and Johnson, C.A. Balancing program and research integrity in community drug abuse prevention: Project STAR approach. *J Sch Health* 56:389-393, 1986.

Pentz, M.A.; Dwyer, J.H.; MacKinnon, D.P.; Flay, B.R.; Hansen, W.B.; Wang, E.Y.I.; and Johnson, C.A. A multi-community trial for primary prevention of adolescent drug abuse: Effects on drug use prevalence. *JAMA* 261:3259-3266, 1989.

Pentz, M.A.; Hansen, W.B.; Rippentrop, K.; Johnson, C.A.; and Flay, B.R. The Midwestern Prevention Project. Formative evaluation of high school booster: Implications for longitudinal drug prevention programming in adolescence, submitted for publication.

Perry, C.L.; Klepp, K.I.; Helper, A.; Hawkins, K.H.; and Murray, D.M. A process evaluation study of peer leaders in health education. *J Sch Health* 58:82-87, 1988.

Perry, C.L., and Murray, D.M. Preventing adolescent drug abuse: Implications from etiological, developmental, behavioral, and environmental models. *J Primary Prev* 6:31-52, 1985.

Raudenbush, S.W., and Bryk, A.S. A hierarchical model for studying school effects. *Sociol Educ* 59:1-17, 1986.

Rogers, E.M. *The Diffusion of Innovations.* New York: Free Press, 1983.

Schaps, E.; Moskowitz, J.M.; Malvin, J.H.; and Schaeffer, G.A. Evaluation of seven school-based prevention programs: A final report on the Napa project. *Int J Addict* 21:1081-1112, 1986.

Sobol, D.F.; Rohrbach, L.A.; Dent, C.W.; Gleason, L.; Brannon, B.R.; Johnson, C.A.; and Flay, B.R. The integrity of smoking prevention curriculum delivery. *Health Educ Res* 4:59-68, 1989.

Sussman, S.; Dent, C.W.; Brannon, B.R.; Glowacz, K.; Gleason, L.R.; Ullery, S.; Hansen, W.B.; Johnson, C.A.; and Flay, B.R. The television, school, and family smoking prevention/cessation project. IV. Controlling for program

success expectancies across experimental and control conditions. *Addict Behav* 14(6):601-610, 1989.

Tricker, R., and Davis, L.G. Implementing drug education in schools: An analysis of the costs and teacher perceptions. *J Sch Health* 58:181-185, 1988.

AUTHORS

Mary Ann Pentz, Ph.D.
Associate Professor

Elizabeth Trebow, M.A.
Research Assistant

Institute for Health Promotion and Disease Prevention Research
Department of Preventive Medicine
University of Southern California
Suite 200
35 North Lake Avenue
Pasadena, CA 91101

1H0143

Experimental and Quasi-Experimental Designs in Prevention Research

David L. Snow and Jacob Kraemer Tebes

INTRODUCTION

A central purpose of social experiments and quasi-experiments is to maximize the possibility of making valid causal inferences. The importance of establishing such cause-effect relationships in the prevention field is to enhance the degree of certainty about which types of interventions (defined in terms of their content, timing, intensity, duration, and other dimensions) have the greatest likelihood of reducing the incidence of maladaptive behaviors in the target populations of interest. Specifying causal linkages makes it possible to identify the most essential elements of a program and to increase its effectiveness with various types of participants (Hormuth et al. 1985).

In this chapter, the authors first discuss briefly the broader context of research design to identify certain predesign issues that unless addressed adequately reduce the usefulness of any research endeavor regardless of what experimental design is employed. Concepts of causal inference and control are then reviewed before outlining the basic types of validity and the threats to validity that can occur. The chapter concludes with a delineation of selected experimental and quasi-experimental designs, along with a discussion of their major advantages and disadvantages, and data analytic techniques.

PREDESIGN ISSUES

The strength of any experimental or quasi-experimental prevention design is dependent on the careful specification of (1) the problem to be prevented, (2) the target population, (3) the risk factors and associated mediating processes that will be the focus of the change effort, (4) the intervention(s) to be employed, and (5) the expected outcomes and related evaluation criteria. These decisions should be based on theoretical and empirical evidence and constitute the first stage in any prevention research project.

The initial steps in planning a preventive intervention are concerned with problem definition and determination of the base rate at which the problem occurs within the population. In substance abuse prevention research, this requires a clear delineation of the different types of substances or classes of substances to be prevented and, for each, a determination of what constitutes problem behavior. For instance, is the behavior classified as a problem on the basis of use versus nonuse or on the basis of a certain level of use defined as abuse? Using these criteria, epidemiological evidence concerning incidence and prevalence rates of substance use or abuse then can be used to specify the target population for intervention.

The selection of risk factors and associated mediating variables that will be the focus of the intervention is based on evidence from risk factor research—the factors and processes that place the population at risk; that is, which individual or situational factors are associated with substance abuse (from cross-sectional risk assessments) and which factors or processes are predictive of substance abuse (from longitudinal risk assessments)? Such evidence serves to specify the factors that, if modified, have some likelihood of leading to a reduction in incidence rates of the problem behavior.

Development of an appropriate and potentially effective intervention involves specifying, in operational terms, the program elements that are viewed as likely to modify the identified risk factors or to effect change in variables that mediate risk status and outcome. Prior conceptual or empirical work is necessary for making choices about the content, timing, intensity, and duration of an intervention. As Cook and Campbell (1979, p. 345) state, "To conduct a randomized experiment with hastily chosen treatments is a waste of resources; one has results about a treatment of little interest or a treatment that had little promise of resulting in effects." Exploratory research and pilot testing can help to ensure the workability of an intervention in a given setting and its operational stability over the course of implementation.

The final predesign issue involves the specification of expected outcomes and the choice of appropriate evaluation criteria. Here, a distinction is drawn between proximal programmatic objectives and distal prevention goals (Heller et al. 1980). Reliable and valid measures need to be selected or developed, therefore, to answer two questions: Did the intervention have an effect on the risk factors or mediating variables targeted for change (proximal objectives)? and Are these changes linked to an ultimate reduction in rates of the end-state variables of interest (distal goals)? Within this framework, specific hypotheses can be generated along with how predicted changes will be detected.

Causal Inference and the Problem of Control

To infer the existence of a cause-effect relationship and to rule out plausible, alternative interpretations, certain criteria need to be met (Cook and Campbell 1979; Mahoney 1978). The first is that covariation be demonstrated between the presumed cause and effect, that is, that manipulation of a cause will result in the manipulation of an effect. The second criterion is the need for relative temporal contiguity between the cause and effect. Third, the effects must follow causes in time, and fourth, all possible influences other than the independent variable(s) in question need to be eliminated or controlled. The fifth criterion is that the causal relationship must be replicable. The first three criteria are givens for prevention intervention researchers. The fourth criterion focuses on issues of internal and statistical conclusion validity. The fifth deals with issues pertaining to external and construct validity.

To infer causal relationships from social experiments and quasi-experiments, it is essential that the investigator effectively address the problem of control. Control is necessary to ensure that observed changes in outcome variables are due to the effects of the intervention as opposed to confounding or extraneous variables. Because special problems of control are encountered in research involving open systems, such as real-world settings, detecting such causal relationships becomes more difficult.

As Higginbotham and colleagues (1988) have summarized, a combination of three diverse models of control typically is employed in research on social and psychological phenomena, each based on different assumptions and containing certain limitations in applicability. The first, derived from the physical sciences, involves *isolation* of the phenomena of interest from any extraneous factors that may influence outcome. Although strenuous efforts are made in prevention research to reduce the possibility of "contamination" effects, the conditions of isolation possible in closed systems are basically impossible to achieve in open systems research. Among the many potential contaminants that occur are the spread of treatment effects to the control group; changes in structural, policy, or procedural characteristics of settings; and the introduction of other programs to study participants.

The second model, derived from agricultural research, involves the *random assignment* of units to treatment conditions. In this model, there is an interest in studying the effects of the treatment under natural conditions. Although randomization can ensure that groups are equivalent before intervention, humans, unlike plants, are active recipients resulting in a certain loss of control. Research participants, for example, will show selective attention, will drop out because of disinterest, will relocate to a different area, or will want a different

treatment. Such changes will lead to some degree of noncomparability of groups at the completion of the intervention.

The third model involves *statistical control* and was developed in social science research in which manipulation of variables, such as demographic characteristics, is not possible or feasible. Statistical procedures are used in attempts to eliminate the influence of extraneous factors that may have a causal relationship to the outcome. The level of control possible is dependent on identifying all those extraneous variables that are related to outcome, reliably measuring them, and identifying the direct and interactive effects of the extraneous variables on the outcome of interest. Difficulties in meeting these conditions result in uncertainty about the success of statistical adjustments in eliminating extraneous influences.

Problems in control introduce plausible, rival explanations to the intervention as the cause of observed changes in the dependent variables. Although combinations of these models can be used to enhance control, some loss of control will always occur. As a result, the investigator must develop a list of rival hypotheses, carefully explore these to reject some, and be left with as few plausible alternatives as possible (Hormuth et al. 1985).

TYPES OF VALIDITY

In a recent reformulation of Campbell and Stanley's (1966) classic monograph on experimental and quasi-experimental designs in field settings, Cook and Campbell (1979) provide the foundation for understanding validity issues in research. They describe four types of validity: statistical conclusion validity, internal validity, construct validity, and external validity. What follows is a brief summary of each of these types of validity as discussed by Cook and Campbell (1979) as well as other investigators such as French and Kaufman (1981), Higginbotham et al. (1988), Hormuth et al. (1985), and Wortman (1983). This section concludes with a discussion of the relative priorities of each type of validity for prevention intervention researchers.

Statistical Conclusion Validity

Statistical conclusion validity refers to the researcher's ability to draw a valid inference about the relationship between cause and effect variables. Cook and Campbell (1979) identify seven significant threats to statistical conclusion validity: inadequate statistical power, violations of assumptions of statistical tests, the use of multiple tests of significance to increase the chance of a type I error, the use of unreliable dependent measures, the existence of random variability in either the intervention setting or the intervention respondents, and

the unreliable implementation of the intervention. To this list, Wortman (1983) has added an eighth major threat to statistical conclusion validity: the occurrence of errors in coding and recording data.

Although each of these threats requires close attention by investigators, two are frequently overlooked by intervention researchers: issues involving statistical power and the reliable implementation of the intervention. In planning studies, it is essential that the investigator determine whether there exists sufficient power to detect an effect of a specific magnitude in a particular sample. In most discussions concerning power, the emphasis is usually placed on having investigators make sure that their sample size is sufficiently large to detect a difference between the intervention and control group. Although absolute sample size is critical, investigators also must keep in mind that power is influenced by the relative sample sizes of the intervention and control groups as well as by misclassification and measurement errors that are present in the study (Rosenbaum 1987). Admittedly, it is often a time-consuming and complex process to conduct a power analysis when planning an intervention study. The information required usually involves obtaining accurate estimates of the effect of the planned intervention for a particular sample, information that is often not readily available. Fortunately, some excellent texts for power analysis that include tables and formulas are available (Cohen 1977; Fleiss 1981; Kraemer and Thiemann 1987).

Another frequent problem for intervention researchers is making sure the intervention is implemented in a reliable manner. Variability in implementing an intervention has the effect of reducing power by inflating error variance and thus making it more difficult to observe a true difference between groups when one actually exists. Instituting a process evaluation along with one that assesses program outcome or impact enables the investigator to make adjustments while the intervention is under way or, at the very least, to identify sources of error that may help explain the findings.

Internal Validity

Internal validity refers to the investigator's ability to determine whether the observed relationship between cause and effect variables may be attributable to the intervention. When such differences cannot be so attributed, they represent "plausible rival hypotheses." In laboratory studies, once the investigator is satisfied that the design provides for considerable experimental control, such rival hypotheses about the effects of the independent on the dependent variable are of minimal concern. In field settings, and particularly in preventive intervention research, almost the opposite is true. Even the best design implemented under ideal conditions does not enable the investigator to be

assured that validity threats will not emerge as the intervention proceeds. For this reason, close attention to potential internal validity threats are at the heart of the prevention researcher's task.

Twelve such threats to internal validity have been identified (Cook and Campbell 1979; Higginbotham et al. 1988; Wortman 1983)—eight involve the potential systematic influence of the dependent variable; three involve changes that affect the integrity of the control and/or intervention group; and one is unique to cross-sectional studies that do not involve an intervention. With the exception of this last threat, each of these is summarized briefly in the discussion that follows.

History, maturation, testing, instrumentation, statistical regression, selection, differential mortality, and interactions with selection that involve maturation, history, and instrumentation are internal validity threats that all involve some systematic influence of the dependent variable. *History* refers to the potential that the observed effect results from a specific event or condition that occurred between the pretest and subsequent measurement. *Maturation* refers to the potential changes in a respondent's growth and experience subsequent to the pretest that may influence the observed effect. *Testing* describes the possible confounding caused by having respondents complete the same or similar tests at pretest and posttest to measure the observed effect. *Instrumentation* refers to the potential threat posed by changes in the test or instrument between measurements. *Statistical regression* describes the phenomenon by which respondents classified into groups on the basis of high or low pretest scores are likely to regress toward the mean in subsequent measurements, particularly if pretest measures were unreliable or contained a significant amount of measurement error. *Selection* refers to the potential for an observed effect to be due to systematic differences between groups that are apparent at pretest. *Differential mortality* describes the threat posed when respondents over time or in different groups differentially fail to complete the study. Finally, *interactions of maturation, history, or instrumentation with selection* that influence the observed effect pose additional threats to validity.

Prevention researchers can minimize these potential threats to internal validity in a variety of ways. First, if at all possible, investigators should use random assignment to intervention groups. All the threats that involve some systematic influence of the dependent variable can be ruled out if this is done. If this is not possible, or if the initial randomization procedure breaks down over the course of the study, investigators should consider a second approach to minimizing such threats to internal validity, that is, making direct adjustments in the study design (Higginbotham et al. 1988). For example, in the absence of random assignment, the effects of maturation may be minimized by using additional

pretest measures to detect a developmental pattern or trend. Similarly, the effects of history can be minimized by ensuring that the intervention is delivered and the pretest and posttest measures are administered during times that do not overlap with those used in subsequent replications. Furthermore, the effects of differential mortality or attrition can be minimized (or at least monitored) by obtaining detailed measures of participants' characteristics at the pretest or, if this is not possible, through exit or followup interviews. As Hormuth and colleagues (1985) have suggested, the choice of just what should be measured is often informed by having the investigator develop a theory of attrition before implementing the study. Careful monitoring of participants who fail to complete the study helps eliminate rival plausible explanations for the obtained findings as well as improves the study's external validity. A third approach to minimizing some of the above threats to internal validity is obtaining multiple pretest observations (Hormuth et al. 1985). For example, multiple pretest measures may identify differences in selection or in interactions with selection that differentiate the groups, or they may reveal problems involving instrumentation, especially when measurements require observations. Finally, a fourth approach to reducing the internal validity threats discussed above is through statistical controls. It is not uncommon for nonequivalent groups to differ on some measures at pretest and to have these measures be systematically related to measures of outcome. When such differences are discovered, they can often be minimized through statistical adjustments such as analysis of covariance. Although such procedures are frowned on by laboratory investigators, it is often the only remaining option for field researchers.

Threats to internal validity also arise from systematic influences of the control and/or intervention groups. The first of these involves *treatment contamination* (Higginbotham et al. 1988). This refers to the potential for intervention and control respondents within the same or proximal settings to communicate with one another or for control respondents to become exposed to the intervention or its equivalent. As Cook and Campbell (1979) point out, treatment contamination can occur either through the "diffusion of treatments" within a setting or through "compensatory equalization of treatments" by well-intentioned personnel who wish to correct the inequity of goods and/or services received between the treatment and control groups. The other type of internal validity threat that reflects systematic changes in the control group involves *atypical responses of control participants* (Higginbotham et al. 1988). This occurs when control respondents learn that they are receiving fewer goods and services than the intervention group and thus begin to adjust their responses accordingly. Some respondents may work harder in trying to overcome this difference while others may essentially give up (Cook and Campbell 1979; Higginbotham et al. 1988), thus introducing bias that may influence the observed effect.

Another related category of threat to internal validity involves *systematic changes in reporting bias* (Higginbotham et al. 1988). This threat involves both intervention and control respondents and focuses on how knowledge of one's group assignment can influence self-reporting. For example, respondents who believe they may have something to gain by overreporting competencies or deficits may do so to receive a desired intervention. Because self-reports completed subsequent to pretest would no longer require such a strategy, the validity of such self-reports will have been tainted by systematic response bias.

Reducing these three internal validity threats often poses more of a problem for the prevention researcher, due to the increased costs involved, to the encountering of real-world limitations, or to both. One obvious strategy is to separate the various intervention and/or control groups of the study. This strategy is often not completely feasible because assignment to groups in preventive intervention studies commonly occurs within a single setting. Nevertheless, investigators usually do have some control over when and where interventions take place, thus making it possible to minimize contact among participants from different groups. As a general rule, investigators should discuss this issue openly with their contact person from the institution in which the intervention will take place so that programmatic differences between groups are not advertised by the institution as part of its effort to recruit participants. A second approach to deal with these internal validity threats is to monitor the intervention closely. In the best of circumstances, this involves implementation of a comprehensive process evaluation along with every preventive intervention. Such an evaluation enables the investigator to identify directly issues of treatment contamination as well as likely reasons for atypical reactions of control participants or systematic changes in reporting bias. The process evaluation should be driven in part by a prior theory of likely participant reactivity and response bias that the investigator can use to identify appropriate items to include in the assessment. When such a comprehensive evaluation is not possible, prevention investigators might consider less costly alternatives such as focus groups with study participants or intervention staff, administration of brief paper-and-pencil surveys accompanying other scheduled events of the study, or individual interviews with key project personnel. Whenever possible, process assessments should be used to provide corrective feedback concerning the three validity threats described above.

Construct Validity

Construct validity involves the conceptualization and operationalization of theoretical constructs for the manipulated and observed variables. At the conceptual level, construct validity refers to the hypothesized causal relationships that the investigator postulates to account for the relationship

between cause and effect variables (Higginbotham et al. 1988). At the operational level, construct validity refers to the translation of these hypothesized relationships into independent and dependent variables.

Threats to construct validity result from inadequate operationalizations of the theoretical independent variable or the use of inaccurate indicators of the theoretical dependent variable. Cook and Campbell (1979) have identified 10 threats to construct validity of causes and effects. Although this list is too exhaustive to be reviewed here, potential threats to construct validity can be divided into three reasonably distinct groups: those that focus on establishing reliable operationalizations of the independent and/or dependent variable (seven in all), those that deal with generalization of the hypothesized construct to other constructs (two in all), and one threat that involves the inadequate conceptualization of hypothesized constructs.

Prevention intervention researchers have several strategies that they can use to improve construct validity. As Campbell and Fiske (1959) have shown in their discussion of the multimethod-multitrait matrix, construct validity can be strengthened by identifying multiple indicators of similar as well as different conceptual (dependent or independent) variables and then assessing whether those that are similar are correlated and those that are different are uncorrelated. In prevention intervention research, this is often relatively easy to do with dependent variables and quite difficult with independent variables. For example, suppose that high-risk youth are the target of a personal and social skills training program that emphasizes peer resistance strategies as a means to reduce experimentation with illicit substances. It is usually easier to include an additional dependent measure, such as involvement with criminal activity (an indicator that can be expected to be correlated with experimentation with illi drugs), than to introduce an additional intervention correlated with peer resistance strategies, such as decisionmaking skills. This additional indicator of the independent variable would require a doubling in the number of participants needed for the study. Another way to strengthen construct validity is to conduct replications of the same experiment in which the construct studied is varied only slightly from one experiment to another. A limitation of this approach is that such multiple replications can be costly and difficult to achieve in field settings. A third approach to improving construct validity is through utilization of structural equation models (Bentler and Newcomb 1986). In this approach, the investigator specifies relationships among the variables in a preventive intervention based on an *a priori* construct model. The accuracy or fit of this conceptual model can then be tested statistically. One significant limitation of this approach is that the sample size required for model testing often exceeds the number of participants available in the intervention setting.

External Validity

External validity deals with the extent to which the causal relationships between theoretical independent variables and theoretical dependent variables are generalizable. Questions concerning generalizability may take either of two forms. First, does this causal relationship generalize *across* different populations, different settings, and different times? Second, does this relationship generalize *to* specific populations, specific settings, and specific times? Typically, basic researchers concern themselves with the former question, while applied researchers are concerned with the latter. Threats to external validity are of three types: interactions of participant selection, setting, and history with the intervention. Threats involving *selection by intervention interactions* usually result from failure to obtain a representative sample from the population of interest. Without a representative sample, the investigator cannot be sure that the observed effect is generalizable across different persons and to a specific population. *Setting by intervention interactions* refers to threats to external validity that arise from having a particular intervention interact with a particular setting in which it was delivered to produce the observed effect. Finally, *time by intervention interactions* refers to external validity threats that result from having the observed effect occur at a particular time in history.

External validity can be strengthened in several ways (Cook and Campbell 1979; Higginbotham et al. 1988). First, random assignment of participants to groups provides the best protection against failure to achieve representativeness. When true representativeness is not possible (and for all intents and purposes, it never is), investigators should consider an alternative strategy of sampling to ensure that groups are heterogeneous. When generalizations are to be made *across* persons, investigators should attempt to sample for heterogeneity; when generalizations are to be made *to* a specific population, investigators should try to make sure that the sample is as representative of that particular population as possible. This enables the investigator to have greater confidence that the preventive effect of an intervention will hold across persons or to similar persons across settings. An obvious means of improving the external validity of preventive interventions is to conduct replications. Unfortunately, few incentives exist for replication studies because of funding limitations, journal policies, and pressures to undertake original investigations. Finally, molecular analyses that attempt to identify relevant interactions also can improve external validity (Higginbotham et al. 1988). Although such analyses frequently restrict the generalizations that can be made about the observed effect to specific persons, settings, or times, they increase one's confidence that the observed effect is applicable under specific boundary conditions.

Priorities Among Validity Types in Prevention Research

In conducting social experiments and quasi-experiments, investigators are usually confronted with the problem that strengthening one type of validity often weakens another. This dilemma requires the investigator to accept tradeoffs among the validity types based on the ultimate objective of the research. Because the primary aim of intervention research is to make causal inferences about the relationship of at least two variables, validity types are usually prioritized according to what best achieves this aim (Cook and Campbell 1979). In most instances, this criterion gives the greatest priority to internal validity for basic and applied researchers. The similarity between these two types of researchers ends here, however.

According to Cook and Campbell (1979), applied researchers generally rank internal validity first, followed by external validity, construct validity of the dependent variable, statistical conclusion validity, and construct validity of the independent variable. Basic researchers, on the other hand, rank construct validity of the independent variable next after internal validity, followed by statistical conclusion validity, construct validity of the dependent variable, and external validity.

The differences between applied and basic researchers is readily apparent when considering the nature of prevention intervention research. Apart from determining whether the intervention causes a change in the dependent variable (an example of internal validity), the prevention investigator usually wants to know in descending order: (1) To whom is the effect applicable (external validity); (2) what, in particular, is affected (construct validity of the dependent variable); (3) was there sufficient experimental control to warrant the conclusion drawn (statistical conclusion validity); and (4) what specific aspects of the intervention caused the observed effect (construct validity of the independent variable). Basic researchers, on the other hand, are usually very interested in identifying key theoretical constructs responsible for the observed effect and are usually least interested in generalizing to specific populations for which the observed effect might apply.

EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS

In this section are six designs that we believe are well suited for evaluating preventive interventions. Space limitations preclude the review of other applicable designs. The criteria we used to select a particular design is that, at least in theory, it be relatively straightforward to implement, require reasonable sample sizes, and provide interpretable results. Three of these are True Experimental Designs—a Pretest/Posttest Control Group Design, an Attention

Placebo Six Group Design, and a Repeated Measures Design; the other three are Quasi-Experimental Designs—a Pretest/Posttest Nonequivalent Control Group Design, a Simple Time Series Design, and a Multiple Time Series Design (which is just a special case of the Nonequivalent Control Group Design). Each design is described separately, along with its major advantages and disadvantages and the common data analytic techniques employed in its application.

The notational system used to describe each design below follows that employed by Campbell and Stanley (1966) in that "R" refers to a random assignment, "O" refers to an observation, "X" refers to the intervention, "Y" refers to the placebo control, and "T" refers to time of testing. Observations located before an intervention represent pretest measures; those located after intervention represent posttest measures.

True Experimental Designs

All the True Experimental Designs described below have the advantage of random assignment of participants to groups. Randomization controls for most of the common threats to internal validity such as history, maturation, instrumentation, statistical regression, selection, testing, differential mortality, and various interactions with selection. However, randomization is also a basic disadvantage when such group assignments are not possible.

Under ideal conditions, in all three experimental designs, testing occurs under "blind" conditions such that the investigator is unaware of which group is being tested at any given time and the measures are reliable and valid as well as sensitive to changes in the dependent variable. The sample employed is of sufficient size to detect a true difference between the groups should one exist (in the case of the Solomon Four Group, Attention Placebo, and Repeated Measures Designs, sample sizes are equal). Finally, for all three designs, the intervention is implemented in a reliable manner.

Pretest/Posttest Control Group Design. In the Pretest/Posttest Control Group Design, participants are randomly drawn and randomly assigned to intervention and control groups; only the intervention group receives the independent variable, and both groups are pretested and posttested on the dependent variable.

R O X O
R O O

This design has distinct advantages. It is simple yet allows for a controlled test of the intervention with relatively few resources required for implementation. The major disadvantage of the Pretest/Posttest Control Group Design is its failure to control for effects due to the interaction of testing and the intervention.

This design can be analyzed using t-tests or analysis of variance applied to posttest scores, analysis of covariance of posttest scores with pretest scores as the covariate, or analysis of covariance through the use of multiple regression.

Attention Placebo Six Group Design. This design represents an improvement over the Solomon Four Group Design. In the Solomon Four Group Design, participants are randomly drawn and randomly assigned to one of four groups, two of which receive the intervention and two of which serve as controls. One intervention group is tested both before and after the intervention, while the other is only tested after. The same testing procedure is also followed for the two control groups.

In the Attention Placebo Six Group Design, participants are randomly drawn and randomly assigned to one of six interventions, two of which receive the intervention, two of which serve as no-intervention controls, and two of which serve as attention placebo controls. Similar to the Solomon Four Group Design, one intervention group is tested both before and after the intervention, while the other is only tested after, with the same procedure being used for the two no-intervention control groups. Both attention placebo control groups receive the placebo intervention, although one is tested both before and after receiving the placebo, while the other is tested only after.

Solomon Four Group Design

R	O	X	O
R	O		O
R		X	O
R			O
R	O	Y	O
R		Y	O

The Attention Placebo Six Group Design is among the most powerful of experimental designs available. It has the advantage over the Solomon Four Group Design in controlling for the effects of expectancy due to participation in an experiment as well as improving on construct validity. Its major disadvantages are the large number of staff resources and participants needed to carry out the design; the increased potential for differential experimental mortality because of the increased number of groups; and the even greater

chance for treatment contamination, atypical responses of control participants, or systematic changes in reporting bias because of the likelihood of proximal groups coming into contact with one another.

This design can be analyzed using a 1 X 3 analysis of variance with interactions, with posttest scores as the dependent variable. One also can use pretest scores as the covariate in an analysis of covariance or conduct an analysis of covariance in a multiple regression format.

Repeated Measures Design. In the Repeated Measures Design, participants are randomly drawn and randomly assigned to one or more intervention and control groups, with only the intervention group(s) receiving the independent variable and all groups being pretested and posttested on the dependent variable. In their most ambitious form, Repeated Measures Designs resemble the Attention Placebo Six Group Designs (with six or even more groups) with multiple measurements taken before and after multiple implementations of the intervention. Such more comprehensive designs also may be factorially organized among categories or groups of participants (e.g., males-females). For the sake of simplicity, we discuss only a simple form of Repeated Measures Design here in which participants are randomly drawn and randomly assigned to one of four groups and each group is measured at three points in time. The first group receives the intervention after time 1 and time 2; the second receives it only after time 1; the third only after time 2; and the fourth group does not receive any intervention.

	T1		T2		T3
R	O	X	O	X	O
R	O	X	O		O
R	O		O	X	O
R	O		O		O

Repeated Measures Designs represent the most powerful of the experimental designs because they maximize each of the four types of validity. Internal and external validity threats are controlled; statistical conclusion validity is enhanced by having participants serve as their own controls; and construct validity of both the independent and dependent variable is strengthened because of the more frequent manipulations of the former and multiple measurements of the latter. The major disadvantages of this design are the considerable resources usually required to carry it out and the very high likelihood of experimental attrition.

Repeated Measures Designs are most commonly analyzed using a multiple analysis of variance and a repeated measures analysis of variance and covariance.

Quasi-Experimental Designs

All quasi-experimental designs have the potential advantage of being particularly well suited to real-world constraints. They allow limited causal inferences even when a randomized experiment has failed. The major disadvantage is the lack of randomization of the sample, which makes causal inferences only as plausible as the comparability of the groups or intervals sampled.

Pretest/Posttest Nonequivalent Control Group Design. In a Pretest/Posttest Nonequivalent Control Group Design, participants are selected for inclusion in either an intervention or control group without random assignment. The remainder of the design resembles the Pretest/Posttest Control Group Design in that only the intervention group receives the independent variable and both groups are pretested and posttested on the dependent variable. In addition, under ideal conditions, testing occurs under "blind" conditions such that the investigator is unaware of which group is being tested at any given time and the measures are reliable and valid as well as sensitive to changes in the dependent variable. Furthermore, the sample employed is of sufficient size to detect a true difference between the groups should one exist, and the intervention is implemented in a reliable manner.

O X O

O O

The Pretest/Posttest Nonequivalent Control Group Design has the advantage of being appropriate for many settings and with populations that may be difficult to assign randomly to groups (e.g., persons at high risk or in imminent danger). This enables investigators in the field to study phenomena that otherwise would not be studied systematically. The major disadvantage of this design is the lack of comparability of the intervention and control groups due to the absence of random assignment, resulting in weakening of threats to internal validity. In addition, when groups are assembled based on extreme scores on a screening measure, there is the additional potential threat posed by statistical regression. With careful monitoring of scores on pretest measures, nearly equivalent groups can be obtained, which allows for reasonable causal inferences to be drawn.

This design can be analyzed using correlated sample t-tests on posttest scores, analysis of covariance of posttest scores with pretest scores as the covariate, or analysis of covariance through the use of multiple regression.

Simple and Multiple Time Series Designs. In the Simple Time Series Design, multiple measurements are taken on a single group before and after the intervention.

O O O X O O O

By comparison, the Multiple Time Series Design is actually a special case of the Pretest/Posttest Nonequivalent Control Group Design in which multiple measurements are taken of two or more nonequivalent groups before and after introduction of the intervention.

O O O X O O O

O O O O O O

Ideally for both time series designs, measurements are unobtrusive or a routine part of the setting so as to minimize the respondents' reactivity to the testing. Measures used are reliable and valid and are sensitive to changes in the dependent variable. In addition, measures are frequent enough in number to be able to detect a linear discontinuity in the measurements taken after introduction of the intervention in the Simple Time Series Design or to detect a linear discontinuity between groups in the Multiple Time Series Design. Finally, the sample size for the two groups is sufficiently large to detect an actual discontinuity should one exist.

Both time series designs have the advantage of being able to provide data retrospectively and unobtrusively, as long as accurate records are available. A further advantage of such designs is the opportunity to obtain additional information about an observed effect through multiple observations. A simple time series also has the advantage of being easy to implement in most social settings. A major disadvantage of the Simple Time Series Design is the absence of a control group (which leaves some obvious validity threats uncontrolled) and the need for relatively large sample sizes.

A Multiple Time Series Design improves on the internal validity of the Simple Time Series Design by providing investigators with a comparison group to assess more accurately the effects of history, maturation, and various interactions with selection. Campbell and Stanley (1966) have described the Multiple Time Series Design as among the best of the quasi-experimental designs because it improves on deficiencies in the Simple Time Series Design and the Nonequivalent Pretest/Posttest Control Group Design. A disadvantage of the Multiple Time Series Design is that external validity threats, such as intervention-testing interactions and intervention-selection interactions, are left uncontrolled, particularly when reactive measurements are employed. Another disadvantage is that the design can require considerable additional resources over and above the Simple Time Series Design and the Nonequivalent Pretest/Posttest Control Group design because of the need for multiple measurements of more than one group.

Cook and Campbell (1979) recommend the use of autoregressive integrated moving average (ARIMA) models and other associated modeling techniques to analyze time series designs. These are preferable to the use of ordinary least squares regression recommended by Campbell and Stanley (1966) because the latter requires that the error terms associated with each time series be independent. When the residuals are independent and the sample sizes are small (50 to 100), repeated measures analysis of variance may be used. When the residuals are correlated, an alternative to using the ARIMA models is analyzing time series data with repeated measures analysis of variance with the Geisser and Greenhouse (1958) correction to degrees of freedom.

CONCLUSION

We have emphasized the advantages of utilizing experimental designs in prevention research. Such designs maximize one's ability to make causal inferences about the effects of an intervention on targeted outcomes because of its greater internal validity. It is essential, however, that prevention intervention researchers who employ experimental designs also be mindful of the importance of attending to issues of external validity. This will enhance the power of generalizations that can be made to specific persons, settings, or times.

Despite our preference for experimental designs, we recognize that there are many limiting factors to a straightforward translation of the experimental model to prevention intervention research, and in some instances, such a model may not be the most desirable. Quasi-experimental designs have been developed in response to these kinds of dilemmas. If attention is given to the soundness and quality of these designs, they allow limited causal inferences to be made. It is

certainly more advantageous in terms of possible knowledge generation to proceed with a well-conceptualized quasi-experimental design than no experiment at all when random assignment is not possible or when a randomized design cannot be maintained. The thoughtful use of both types of designs increases the range of possibilities for meaningful investigations in prevention intervention research.

REFERENCES

Bentler, P.M., and Newcomb, M.D. Personality, sexual behavior, and drug use revealed through latent variable methods. *Clin Psychol Rev* 6:363-385, 1986.

Campbell, D.T., and Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 56:81-105, 1959.

Campbell, D.T., and Stanley, J.C. *Experimental and Quasi-Experimental Designs for Research.* Boston, MA: Houghton Mifflin Company, 1966.

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences.* New York: Academic Press, 1977.

Cook, T.D., and Campbell, D.T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally, 1979.

Fleiss, J.L. *Statistical Methods for Rates and Proportions.* New York: John Wiley, 1981.

French, J.F., and Kaufman, N.J. *Handbook for Prevention Evaluation: Prevention Evaluation Guidelines.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)81-1145. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1981.

Geisser, S., and Greenhouse, F.W. An extension of Box's results on the use of F-distribution in multivariate analysis. *Ann Math Stat* 29:885-891, 1958.

Heller, K.; Price, R.H.; and Sher, K.J. Research and evaluation in primary prevention: Issues and guidelines. In: Price, R.H.; Ketterer, R.F.; Bader, B.C.; and Monahan, J., eds. *Prevention in Mental Health: Research, Policy, and Practice.* Vol. 1. Beverly Hills: Sage Publications, 1980. pp. 285-313.

Higginbotham, H.N.; West, S.G.; and Forsyth, D.R. *Psychotherapy and Behavior Change: Social, Cultural, and Methodological Perspectives.* New York: Pergamon Press, 1988.

Hormuth, S.E.; Fitzgerald, N.M.; and Cook, T.D. Quasi-experimental methods for community-based research. In: Susskind, E.C., and Klein, D.C., eds. *Community Research: Methods, Paradigms, and Applications.* New York: Praeger, 1985. pp. 206-249.

Kraemer, H.C., and Thiemann, C. *How Many Subjects?* Beverly Hills, CA: Sage Publications, 1987.

Mahoney, M.J. Experimental methods and outcome evaluation. *J Consult Clin Psychol* 46:660-672, 1978.

Rosenbaum, P.R. A nontechnical introduction to statistical power and the control of bias. In: Steinberg, J.A., and Silverman, M.M., eds. *Preventing Mental Disorders: A Research Perspective.* National Institute on Drug Abuse. DHHS Pub. No. (ADM)87-1492. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1987. pp. 174-185.

Wortman, P.M. Evaluation research: A methodological perspective. *Ann Rev Psychol* 34:223-260, 1983.

AUTHORS

David L. Snow, Ph.D.
Associate Professor

Jacob Kraemer Tebes, Ph.D.
Assistant Professor

Department of Psychiatry
Yale University School of Medicine
19 Howe Street
New Haven, CT 06511

1H0144

Modeling of Intervention Effects

Peter M. Bentler

INTRODUCTION

An ideal intervention study would utilize an experimental design, randomly assigning individuals to prevention or treatment conditions. The major advantage of such a design is that randomization, if implemented successfully, ensures that extraneous variables cannot provide effective competing explanations for any results that might be observed. In addition, standardly available statistical procedures, such as multivariate analysis of variance (MANOVA) or analysis of variance (ANOVA), are available to analyze the data to yield conclusions about the effectiveness of the intervention. Finally, the appropriateness of the statistical analyses can be evaluated by standard means (e.g., a homogeneity of variance assumption can be checked by well-known methods). As a consequence of these considerations, causal inferences regarding the intervention effects are easy to make and justify.

In practice, the ideal design seems to be difficult to carry out. In the first place, assignment may be carried out at the group level (e.g., classroom) rather than at the individual level. It follows that there may be a lack of independence of observations within a treatment, so the assumptions of standard statistical methods break down and MANOVA and ANOVA no longer have their optimal statistical properties. Second, even if assignment is at the level of the individual, randomization may fail in practice. For example, there may be differential participation rates at assignment, leading to differences between groups at pretest, or there may be differential dropout. Similarly, sample size may be so small that randomization or random sampling will be insufficient to generate equivalence of groups (Hsu 1989). Third, for ethical or other practical reasons, assignment to conditions may depend on the pretest characteristics of the individuals (e.g., the most needy may be assigned the presumed most effective treatment). As a result, statistical control must be used in place of experimental control. The standard method for doing this, analysis of covariance (ANCOVA), may fail to give an unbiased treatment effect because if the control variables are fallible, the control is at the level of an observed variable rather than at the level of the true characteristic. Finally, the experimental conditions may be compromised very badly, for whatever reason,

and some type of within-condition analysis may be needed to salvage any results from a study.

Structural equation modeling cannot solve all of the problems of experimental failure, but it provides one of the most effective currently available methods for imposing theory-based statistical control to substitute for experimental control. At this time, no structural modeling methods are available to deal with the lack of independence of observations within a condition, although research on this topic is under way (Weng and Bentler 1987). Four points are discussed in this chapter: isolating true effects from observed effects, differentiating pretest differences from treatment effects, the role of modeling in the control for missing data, and evaluating effectiveness of treatment using indicators of the degree of program participation.

TRUE VS. OBSERVED EFFECTS

One contribution of structural modeling to the analysis of intervention effects is that, when a model involving latent variables is specified and verified, it allows for differentiating observed from true effects. It has long been known that true effects can be different from observed effects not only in magnitude but also in sign (Lord 1960; Cochran 1968). However, procedures for isolating true effects were considered to depend on estimates of reliability or error variance and were typically made from extraexperimental information, until Sörbom (1978) clarified that an appropriate latent variable design could effectively isolate the relevant effects of interest.

For example, well-known quasi-experimental Head Start study addressed whether children who received a special Head Start educational program subsequently performed better on cognitive tasks than control children who did not receive this extra training. A particular Head Start data set was studied by Bentler and Woodward (1978) and Sörbom (1982). In this study, there were two samples of children. One sample had been given a Head Start educational program; the other sample was a matched control group. Measures on six variables (Vs) were available for both groups. V1 through V4 represent mother's education, father's education, father's occupation, and family income, respectively. They are background variables, essentially to be controlled. The more important variables are V5 and V6, namely, scores on the Metropolitan Readiness Test and the Illinois Test of Psycholinguistic Abilities, respectively. These are the key outcome variables. If the Head Start program were effective and no controls were needed, one could simply evaluate the group's performance on V5 and V6. The means are as follows:

	V5	V6
Head Start Group	19.672	9.562
Control Group	20.415	10.070

It is apparent that the controls have the higher performance scores, even though they did not receive the special Head Start training program. Thus, the Head Start program would immediately be judged a failure. However, what about the background of the two groups of children? The variables measuring socioeconomic status (SES) showed the following pattern of means:

	V1	V2	V3	V4
Head Start Group	3.520	3.081	2.088	5.358
Control Group	3.839	3.290	2.600	6.435

It is apparent that the controls are higher on all SES indicators. Such status certainly preceded the experiment. It is clear that the study could not have been an experiment, because if children had been assigned randomly to conditions, the means on the control variables V1 through V4 should be approximately the same for both groups. Obviously, the program was administered to the children from lower SES families and any matching of pretest characteristics that was done was ineffective. This is perhaps understandable because these families were the most needy. But then, how would one interpret the differences on V5 and V6? Perhaps the Head Start children are disadvantaged by SES because SES influences the types of educational programs generally available, which in turn could influence intellectual learning and, consequently, performance. Thus, one might expect that Head Start children would have been lower on V5 and V6 before the start of the study. It would even be possible that they improved significantly more than the controls as a result of their program but that this effect was masked by large preexisting differences between the groups, with the controls being higher on intellectual performance by virtue of having higher SES.

A standard approach to analysis would be to do an ANCOVA, controlling for V1 through V4 and evaluating the outcome on V5, and perhaps separately for V6, or doing this jointly in a partial MANOVA. But this approach does not take into account the fallible measures of SES available, and it ignores knowledge that may be imposed on intellectual variables. As the discussion makes clear, it is difficult to do anything meaningful in the way of data analysis without having some hypothesis about what processes are at work in these data. Structural modeling requires an explicit hypothesis about the controls needed and also

about the outcome measures. A model for these data suggested by Bentler and Woodward (1978) and studied more thoroughly by Sörbom (1982) is shown in figure 1. This figure and the surrounding analyses are taken from Bentler (1989).

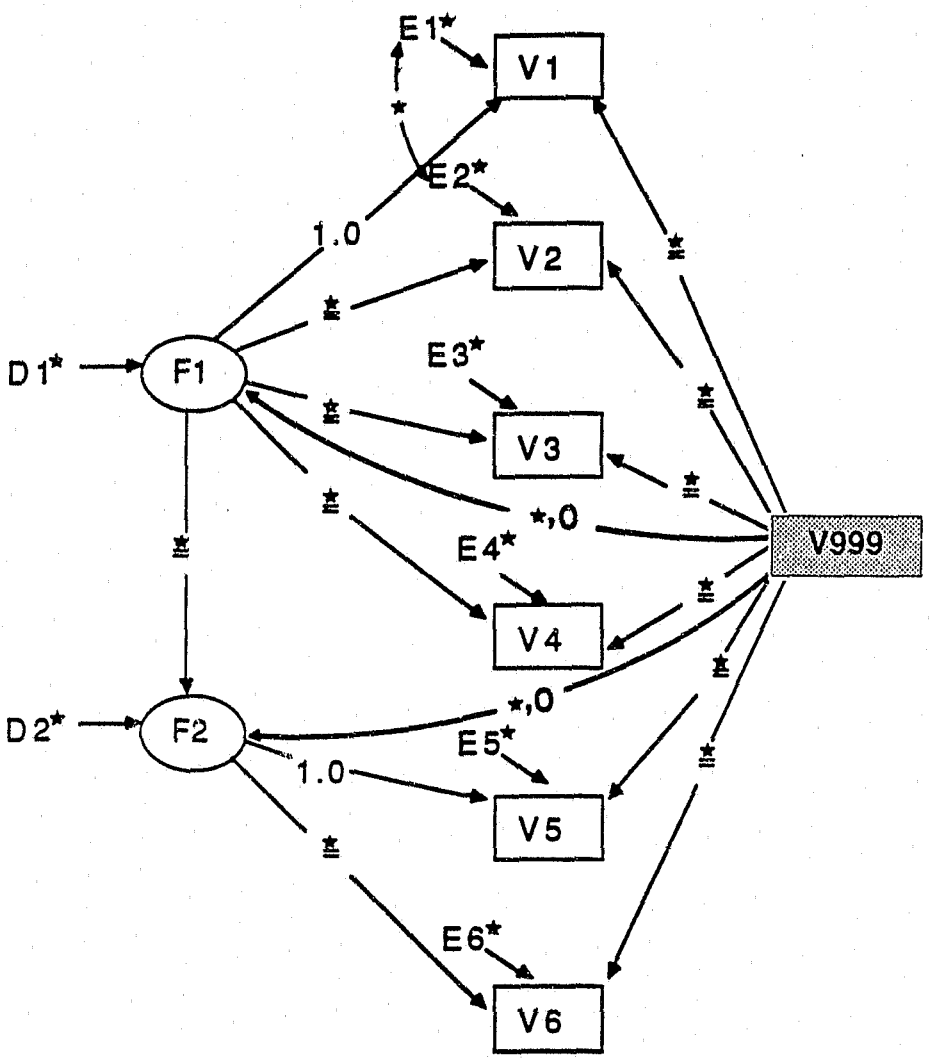

FIGURE 1. *Head Start model*

SOURCE: Bentler 1989, copyright 1989, P.M. Bentler.

The figure uses a standard convention of placing measured variables into rectangles and latent variables or factors into ovals, with directional arrows representing regression coefficients and two-way arrows representing correlations or covariances. The left part of the figure shows two factors, with the factor regression F1→F2. The measured variables are in the middle of the figure. Variables 1 through 4 are indicators of F1, a latent SES factor. Variables 5 and 6 are indicators of F2, a hypothesized ability factor. Each variable has an error residual, with E1 and E2 being correlated. This is a standard factor analysis type of model with two factors, except that the ability factor F2 is regressed on F1, the background SES factor, indicating the hypothesis that SES might affect ability.

The constant "variable" 1.0, designated in the right part of the figure as V999, affects the V variables and the F factors. In this figure, the regression on a constant is an intercept, so the directional arrows from V999 represent such intercepts. In this model, intercepts are hypothesized for all variables, V1 through V6, as well as for both factors, F1 and F2. The intercepts for the V variables are not particularly interesting in the model, because they represent background "levels" for the variables that are common in both groups. The intercepts for the F factors are of special interest. The intercept for the F1 factor (i.e., the V999→F1 path) represents the mean of factor F1 (i.e., the mean of the SES factor). One would expect this mean to be higher for controls than for Head Start participants. The intercept for F2 (i.e., the V999→F2 path) represents the *increment* in means for F2, once the mean on F1 has been controlled (i.e., it is the experimental treatment effect of interest). It would be nice if the Head Start group were higher on this intercept, because it represents (1) an intercept on the latent ability factor that is broader than either single ability indicator V5 or V6 and (2) the effect of the program, given that statistical control for true SES (i.e., F1) has been undertaken. Before discussing the results, a few other points about the figure need to be explained.

The figure, which is not a conventional path diagram, attempts to show the two models of the Head Start and control samples simultaneously. A key to distinguishing the two groups is as follows:

*	Denotes a parameter that is free in both groups
$\overset{*}{=}$	Denotes a parameter that is free in both groups, but constrained to be equal in the two groups
*,0	Denotes a parameter that is free in one group (Head Start) and zero in the other (control) group

Thus, the path diagram makes clear that:

- The variances of E and D variables are free parameters in each group and not constrained to be equal across groups.
- The covariance of E1 and E2 is a free parameter in each group and not constrained to be equal across groups.
- The free factor loadings of all variables (F→V paths) are free parameters in each group, but each loading is constrained to be equal in the two groups.
- The fixed 1.0 factor loadings are fixed for identification at 1.0 in both groups.
- The factor regression (F1→F2) is a free parameter in each group, but the value is constrained to be equal across the two groups.
- The intercepts of the measured V variables are free parameters in each group, but each intercept is constrained to be equal in the two groups.
- The intercepts of the F factors are free parameters in the Head Start group but are set to a fixed zero value in the control group.

Some of these points require further discussion.

Regarding the left part of the figure: If the same factor model holds for both samples, then it is helpful to have the factor loadings and factor regressions be equal across groups. Thus, they are so specified. Equality of residual variances and covariances is not so important and is not imposed.

In the right part of the figure, the paths from V999 to the Vs, each of which is held to be equal across groups, represent a kind of baseline level for the variables. Each path from the intercept V999 to one of the Vs is a "direct" effect. Differences in means of the variables across groups, if they exist, must arise from other sources. In the diagram, these sources can be traced back to the intercepts of F1 and F2. For example, the path V999→F1→V4 makes clear that the intercept of F1 will affect V4, and the path V999→F1→F2→V6 shows that it also will affect V6. These paths are called "indirect" effects. The final means of the V variables are so-called "total effects," which are the sum of the direct and indirect effects. The paths V999→F1 and V999→F2, the intercepts for F1 and F2, reflect coefficients that are free to be estimated in one group but held to zero in the other group (this is done for identification purposes). Consequently, if the freely estimated coefficients are not zero and large, the

differences in factor intercepts will be reflected in mean differences of the observed variables across groups. In this model, the Head Start factor intercepts are estimated, but the control intercepts are set to zero.

Maximum likelihood estimation of the parameters of the model, based on the means and covariance matrices of the Head Start and control samples, using the program EQS, yielded the following results (Bentler 1989, chapter 9). The Head Start equations are as follows:

Measurement Equations With Standard Errors and Test Statistics

V1 = V1 =	1.000 F1	+	3.869*V999	+	1.000 E1
			.094		
			41.084		
V2 = V2 =	.851*F1	+	3.339*V999	+	1.000 E2
	.144		.083		
	5.924		40.314		
V3 = V3 =	1.207*F1	+	2.573*V999	+	1.000 E3
	.222		.090		
	5.430		28.643		
V4 = V4 =	2.758*F1	+	6.421*V999	+	1.000 E4
	.517		.229		
	5.334		28.095		
V5 = V5 =	1.000 F2	+	20.357*V999	+	1.000 E5
			.287		
			70.885		
V6 = V6 =	.850*F2	+	10.085*V999	+	1.000 E6
	.141		.217		
	6.018		46.442		

The measurement equations expressing the relation of measured variables V to the factors F1 and F2, the residual Es, and the intercept V999 are not particularly remarkable. All of the factor loadings (F→V paths) are significant. The variable intercepts are significant, but they just indicate a general level for the variables that is not interpretively interesting. The construct equations for the F variables yielded the following results:

Construct Equations With Standard Errors and Test Statistics

F1 = F1 =	-.382*V999	+	1.000 D1				
	.104						
	-3.685						
F2 = F2 =	2.137*F1	+	184*V999	+	1.000 D2		
	.551		.378				
	3.876		.487				

The regression of F2 on F1, ability on SES, is significant (and equal in both groups; see below). Higher SES children do better than lower SES children. Also, Head Start children were lower in SES to begin with, when -.382 is compared to the control children's value of fixed zero.

The major point of the analysis lies in the V999→F2 path. Note that the experimental Head Start program produced a positive impact (.184) on ability, though the effect is not significant by z-test. This positive impact stands in contrast to the raw variable means presented above, where the controls had the higher means on the ability indicators V5 and V6, and to the total effects of V999 on V5 and V6, shown below, which also verify that the controls have the higher expected variable means in the model. Unfortunately for a judgment of the impact of the program, the V999→F2 path is not statistically significant.

The estimates of variances of the E and D variables are not shown here for either group. These variances were not constrained to be equal across groups, though a more restricted model that imposes such a constraint also could have been considered. Similarly, the estimated covariance of E1 and E2 is not shown for either group. The effect decomposition, in each V999→V path, gives the final estimated mean of each variable, under the model. These are not interesting, except to verify that they are consistent with the sample means, when compared with the control group (shown subsequently).

Decomposition of Effects With Nonstandardized Values
Parameter Total Effects

V1 = V1	=	1.000 F1	+	3.487*V999	+	1.000 E1	+	1.000 D1						
V2 = V2	=	.851*F1	+	3.014*V999	+	1.000 E2	+	.851 D1						
V3 = V3	=	1.207*F1	+	2.112*V999	+	1.000 E3	+	1.207 D1						
V4 = V4	=	2.758*F1	+	5.367*V999	+	1.000 E4	+	2.758 D1						
V5 = V5	=	2.137 F1	+	1.000 F2	+	19.724*V999	+	1.000 E5	+	2.137 D1	+	1.000 D2		
V6 = V6	=	1.817 F1	+	.850*F2	+	9.548*V999	+	1.000 E6	+	1.817 D1	+	.850 D2		
F1 = F1	=	-.382*V999	+	1.000 D1										
F2 = F2	=	2.137*F1	+	-.632*V999	+	2.137 D1	+	1.000 D2						

The estimated mean of a factor is given by the V999→F effect. The estimated mean of the ability factor F2 under the model is -.632 lower for the Head Start children compared with the controls, but this is due basically to the differential SES of the children (i.e., the indirect effect of V999 on F2), which is -.817, statistically significant with z-test value of -2.33.

In the control group, the corresponding results are as follows:

Measurement Equations With Standard Errors and Test Statistics

$$V1 = V1 = 1.000\ F1 + \underset{\substack{.094 \\ 41.084}}{3.869*V999} + 1.000\ E1$$

Additional measurement equations are not shown because all measurement equations have identical estimates and standard errors as in the Head Start group.

Construct Equations With Standard Errors and Test Statistics

$$F1 = F1 = 1.000\ D1$$
$$F2 = F2 = \underset{\substack{.551 \\ 3.876}}{2.137*F1} + 1.000\ D2$$

Of course, the estimated effect of SES on ability is the same as in the Head Start group, by virtue of the constraints imposed.

Decomposition of Effects With Nonstandardized Values
Parameter Total Effects

V1 = V1 = 1.000 F1 + 3.869*V999 + 1.000 E1 + 1.000 D1
V2 = V2 = .851*F1 + 3.339*V999 + 1.000 E2 + .851 D1
V3 = V3 = 1.207*F1 + 2.573*V999 + 1.000 E3 + 1.207 D1
V4 = V4 = 2.758*F1 + 6.421*V999 + 1.000 E4 + 2.758 D1
V5 = V5 = 2.137 F1 + 1.000 F2 + 20.357*V999 + 1.000 E5 + 2.137 D1 + 1.000 D2
V6 = V6 = 1.817 F1 + .850*F2 + 10.085*V999 + 1.000 E6 + 1.817 D1 + .850 D2
F1 = F1 = 1.000 D1
F2 = F2 = 2.137*F1 + 2.137 D1 + 1.000 D2

As can be seen, the model predicts higher means on all measured variables for the controls compared with the Head Start subjects. In essence, this says the model is consistent with the data. In fact, the χ^2 for the model is 27.45, which, with 23 degrees of freedom, indicates the model is statistically acceptable.

PRETEST DIFFERENCES VS. TREATMENT EFFECTS

Another illustration of how structural modeling can help analyze intervention data involves that of an experiment that may have gone wrong, a rather standard occurrence in intervention research. The example is a modest illustrative one, however, again taken from Bentler (1989). Sörbom (1978) reported on an experiment by Olsson on the effects of training on abilities to perform verbal tasks. In a pretest, 11-year-old children were assessed for their verbal ability with two kinds of verbal material, synonyms and opposites. Thereafter, they were randomly assigned to experimental and control conditions. The experimental group received training on similar materials, while the control group did not. Both groups were then retested, yielding posttest data. The pretest and posttest data are to be analyzed.

Sörbom studied several models for these data. A path diagram for his final model is given in figure 2, where V1 and V2 represent the pretest scores on synonyms and opposites tasks, respectively, and V3 and V4 represent synonyms and opposites task performance after the experimental intervention. As seen in the left part of the diagram, it is hypothesized that synonyms and opposites, at each time point, can be conceived as indicators of a latent factor, say, verbal ability. Ability at posttest, F2, is expected to be a function of ability at the pretest, F1. In addition, the residual in opposites at the two time points, E2 and E4, are expected to be correlated from pretest to posttest.

The right part of the figure gives the constant V999. This is presumed to affect each of the variables V1 through V4, reflecting their intercept, a general level for these variables. V999 also directly affects the factors F1 and F2, representing an intercept for these factors. The figure maintains the convention of figure 1 that "*" represents a parameter that is free to be estimated in each group, without constraints; that "*" represents a parameter that is freely estimated in each group but constrained to be equal in both groups; and that "*,0" represents a parameter that is fixed at zero in the control group and is free to be estimated in the experimental group. Thus, factor loadings, factor regressions, and variable intercepts are constrained to be equal across groups. Factor intercepts are fixed at zero for identification in the control group and are free to be estimated in the experimental group. All residual variances and the covariance of E2 and E4 are free to be estimated in each group, without any constraints. The setup for running this model in EQS is given in appendix 1.

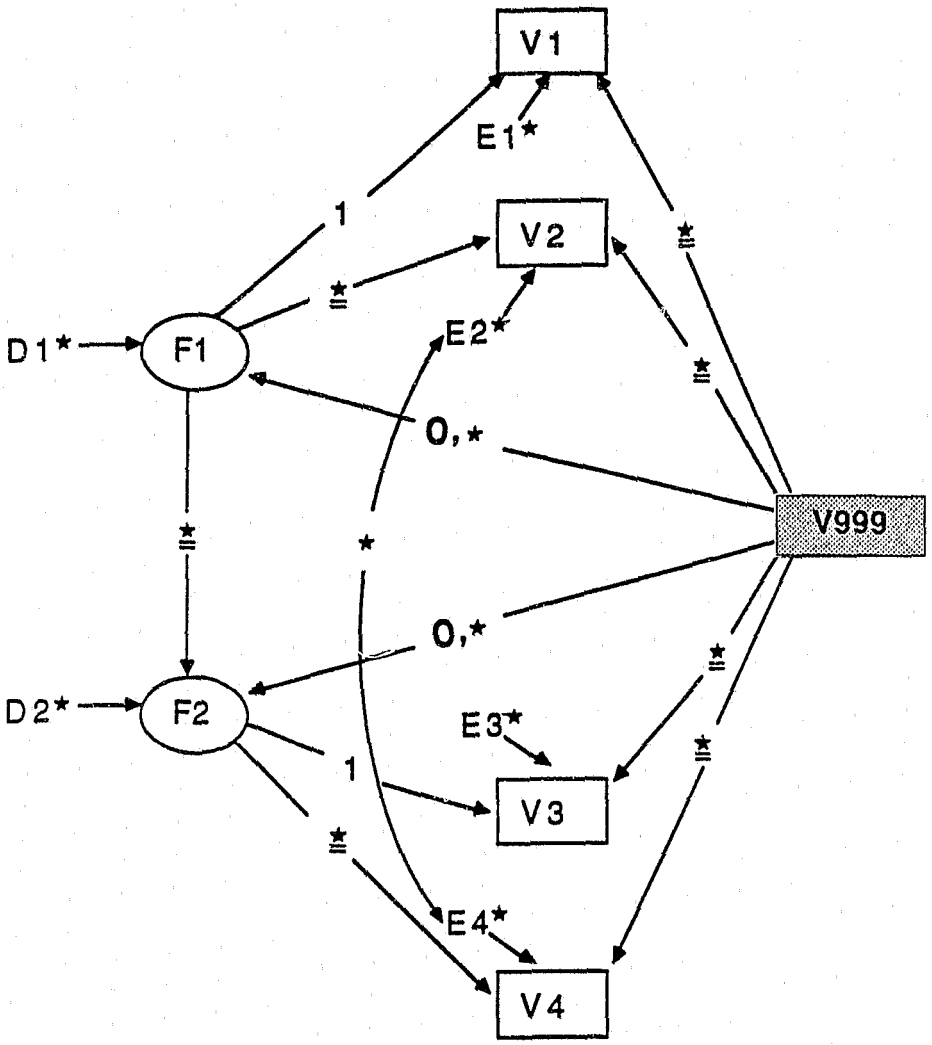

FIGURE 2. *Experimental/control group model*

SOURCE: Bentler 1989, copyright 1989, P.M. Bentler.

As in the previous study, we may examine the means of the observed variables,

	V1	V2	V3	V4
Controls	18.381	20.229	20.400	21.343
Experimentals	20.556	21.241	25.667	25.870

where V1 and V2 are pretest variables and V3 and V4 are posttest. As far as the posttest data are concerned, the experimentals are clearly superior to the controls, so there seems to be good evidence for the experimental manipulation. But the controls seem to be a bit lower at pretest, especially on V1. A look at the covariance matrices of the two groups, shown in appendix 1 in the EQS job file, verifies the additional disturbing fact that the variances of all variables are substantially lower in the control group compared with the experimental group, both at pretest and posttest. There would seem to be some question in these data as to whether the two groups were assigned randomly to conditions or, if so, whether randomization may have failed. In particular, one would expect the pretest means and variances and covariances for the two groups to be equal, but this does not appear to be so. Such a hypothesis could be tested by a structural model but is not done here.

Some results of Sörbom's model (1978) are presented below, first for the controls:

Measurement Equations With Standard Errors and Test Statistics

SYNONYM1	= V1	=	1.000 F1	+	18.619*V999	+	1.000 E1
					.597		
					31.205		
OPPOSIT1	= V2	=	.878*F1	+	19.910*V999	+	1.000 E2
			.051		.544		
			17.286		36.603		
SYNONYM2	= V3	=	1.000 F2	+	20.383*V999	+	1.000 E3
					.538		
					37.882		
OPPOSIT2	= V4	=	.907*F2	+	21.203*V999	+	1.000 E4
			.053		.534		
			17.301		39.719		

The factor loadings look good, and the intercepts for the Vs are about the magnitude of the means in the control group.

Construct Equations With Standard Errors and Test Statistics

ABILITY1 = F1 =	1.000 D1			
ABILITY2 = F2 =	.895*F1	+	1.000 D2	
	.052			
	17.145			

The factor is quite stable from pretest to posttest. The variances and covariances are not presented (to save space), though it should be noted that the variances for D1 and D2 are substantially higher in the experimental compared with the control group. In the experimental group, the following construct equations are obtained (measurement equations are same as for controls).

Construct Equations With Standard Errors and Test Statistics

ABILITY1 = F1 =	1.875*V999	+	1.000 D1		
	.899				
	2.085				
ABILITY2 = F2 =	.895*F1	+	3.628*V999	+	1.000 D2
	.052		.480		
	17.145		7.558		

Because the controls' V999→F1 path is set to zero, the experimentals' comparable path shows that experimental subjects were significantly higher in the verbal ability factor F1 at pretest (z=2.085). Thus, there is some reason to doubt that the two groups were initially equal in ability (the intercept for F1 is the mean, as there are no indirect paths from V999 to F1). Given that they may have been higher in ability, nonetheless the experimentals' training on the verbal materials improved that group's subsequent performance when compared with the controls. This can be seen in the intercept for F2, which is significantly greater than zero, thus reflecting the observed mean differences between the groups on V3 and V4. Overall, the model is also acceptable, with χ^2=3.952, based on 5 df, having an associated probability of .556.

An obvious question is how this model might perform when the intercepts for F1 are forced to be equal for controls and experimentals. Such a specification would be consistent with a randomized assignment of subjects to conditions. Rerunning the above model with the path V999→F1 set to zero in both groups results in a model that is statistically acceptable by the χ^2 goodness-of-fit test. However, one of EQS's diagnostic tests, the Lagrange Multiplier test to evaluate whether the cross-group equality constraints are reasonable, shows that the equality constraint for the two intercepts of V1 across groups is likely to be implausible. Thus, in the next model, this constraint is released, with the following results. The controls' modeling results are given first.

Measurement Equations With Standard Errors and Test Statistics

SYNONYM1	= V1	=	1.000 F1	+	18.738*V999	+	1.000 E1
					.541		
					34.637		
OPPOSIT1	= V2	=	.888*F1	+	20.651*V999	+	1.000 E2
			.051		.442		
			17.353		46.736		
SYNONYM2	= V3	=	1.000 F2	+	20.768*V999	+	1.000 E3
					.465		
					44.659		
OPPOSIT2	= V4	=	.891*F2	+	21.607*V999	+	1.000 E4
			.055		.461		
			16.307		46.880		

Construct Equations With Standard Errors and Test Statistics

ABILITY1 = F1 =	1.000 D1			
ABILITY2 = F2 =	.906*F1	+	1.000 D2	
	.053			
	17.165			

The experimentals' equations follow next.

Measurement Equations With Standard Errors and Test Statistics

SYNONYM1	= V1	=	1.000 F1	+ 20.002*V999	+ 1.000 E1	
				.515		
				38.811		
OPPOSIT1	= V2	=	.888*F1	+ 20.651*V999	+ 1.000 E2	
			.051	.442		
			17.353	46.736		
SYNONYM2	= V3	=	1.000 F2	+ 20.768*V999	+ 1.000 E3	
			.465			
			44.659			
OPPOSIT2	= V4	=	.891*F2	+ 21.607*V999	+ 1.000 E4	
			.055	.461		
			16.307	46.880		

Construct Equations With Standard Errors and Test Statistics

ABILITY1 = F1 =	1.000 D1			
ABILITY2 = F2 =	.906*F1	+ 4.342*V999	+	1.000 D2
	.053	.538		
	17.165	8.075		

The model is acceptable, with χ^2=2.962, based on 5 degrees of freedom, showing a superb fit.

Interpretively, the equal intercepts for F1 in this model across groups suggests that the children in the two groups may well have been equal in mean verbal ability at pretest but that, for reasons that cannot be ascertained from within the analysis, the controls had a lower mean on V1 as well as lower variance on D1 and, hence, on verbal ability F1 (variances not shown above). Isolating the differences between groups in this way allows for an unfettered interpretation of the experimental effect, given as the intercept associated with the path V999→F2. Taken as zero in the control group, the effect in the experimental group is estimated at 4.342 with a standard error of .538, highly significant compared to zero. Thus, the latent variable analysis confirms the observed mean differences in posttest in this case.

This particular model fits even better than the model proposed by Sörbom (1978), but a further understanding of the experimental procedure would be called for to explain the pretest imbalances between conditions.

MODELING IN CONTROL FOR MISSING DATA

Virtually all prevention/treatment studies have a serious problem of attrition. As a consequence, data will be missing and, in general, the longer the span of the study, the greater amounts of missing data. Traditional approaches to analysis of results with missing data involve the use of so-called "listwise deletion," in which a case is eliminated completely if data are missing, or "pairwise deletion," in which a case is eliminated in the computation of summary statistics such as means or correlations if the corresponding data are unavailable. These procedures are practical but problematic. Both are inefficient, that is, do not produce the best possible or least variance estimates. Listwise deletion is also inefficient in another sense: It throws away a substantial amount of potentially useful data. Pairwise deletion sometimes yields correlation matrices that have inappropriate properties (specifically, that are not positive definite).

In recent years, some rather general approaches to missing data have been developed. Two variants of these general approaches (Allison 1987; Muthén et al. 1987) suggest that structural modeling may play a useful role in some situations. In particular, if the missing data contain a few predominant patterns of missing data (e.g., some subjects have data only from waves 1 and 2; all others have complete data from waves 1, 2, and 3), modeling is an attractive approach.

One can distinguish among several interrelated concepts in this literature, going back to Rubin (1976).

> Data are missing at random [MAR] if the probability of obtaining the particular pattern of missing data found in the sample does not depend on the values of the data that are missing. It may, however, depend on the values of the data that are observed. Data are observed at random [OAR] if the probability of obtaining the missing data pattern found in the sample does not depend on the data that are observed; however, it may depend on the data that are missing (Allison 1987, p. 76).

If both of these conditions, MAR and OAR, are satisfied, the data are said to be missing completely at random (MCAR). Listwise and pairwise deletion can be fully justified only when the data are MCAR, which is a very strong assumption not likely to be met in practice.

Allison (1987) and Muthén et al. (1987) have shown that data do not need to be MCAR for structural modeling to yield consistent estimates of the parameters, appropriate standard error estimates, and by some manipulation, appropriate χ^2 tests. In particular, when the data are only MAR, but not necessarily OAR, under mild conditions the procedure produces appropriate inferences. However, even the weaker assumption that the data are MAR may be violated in longitudinal research, because attrition may depend on the values of the variables that would have been observed in later waves. Nonetheless, it appears that the structural modeling approach will more likely yield appropriate (less biased) inferences than listwise or pairwise deletion even when MAR does not hold.

The typical application of structural modeling requires raw scores for all subjects, or means and covariances. Because not all subjects have all data, how can structural modeling proceed? A structural modeling approach to missing data creates groups or samples of subjects in accord with their pattern of missing data. If there are three patterns of missing data, a three-group structural model is used. If there are dozens of patterns of missing data, with only a few subjects showing a given pattern of missing data, this approach is useless because some of these samples may be too small to yield stable results (or a positive definite covariance matrix for observed data) and the method may be too computationally demanding to work with so many samples. In practice, dummy variables and factors are used in the groups with missing data, with pseudovalues replacing the missing means and covariances. Equality constraints across groups are used to ensure that the same parameters (means and covariances, or structural modeling parameters) are estimated in both groups when these parameters would be identified if the data were complete (they may not be identified in any single sample), and the process is carried out so that the pseudovalues are fitted exactly. Both means and covariances must be modeled.

An example of the structural modeling approach to missing data is given in the job setup shown in appendix 2. The data and model are taken from Allison (1987), who used the LISREL program to estimate and test the model; this required using a number of "tricks," such as using dummy variables and parameters and adjusting the degrees of freedom to yield the correct missing data results. These tricks are completely unnecessary to the theory involved and serve to confuse the simplicity of the ideas. They also are not necessary in EQS, in which the model setup is essentially the same as in any multisample analysis with structured means. The critical point in such a setup is that the samples with missing data can contain specifications of equations, intercepts, variances, and covariances only for variables that are actually measured, as well as for hypothesized factors and residual variables relevant to those variables.

The model under consideration is a two-group model in which one sample has complete information on all variables and the other sample containing a particular pattern of missing data, thus containing observed data on a subset of the variables. The data originally came from Bielby et al. (1977), and the specific meaning of the variables in the example can be found in Allison (1987). What is relevant here is that the model is essentially that shown in figure 2, except that (1) the paths from V999 to F1 and F2 are removed from the model, so that the Fs are independent variables and there are no Ds; (2) the path from F1 to F2 is replaced by a two-way arrow, a covariance; and (3) there is no error covariance E2,E4. Thus, a two-group model similar to figure 2 is being evaluated. As seen in lines 6 through 15 of appendix 2, the model is a simple two-factor model with intercepts for the measured variables (paths from V999 to the measured variables). The data from this sample, based on 348 cases, are complete; that is, all variances and covariances among V1 through V4, as well as the means of these variables, are available for analysis, as shown in lines 16 through 22. Thus, in this first group, the model is a rather standard factor analysis model, except that the variable intercepts (means in this example) also are being estimated. If this sample were the only one being analyzed, these intercepts would be estimated at the sample means.

The second, much larger sample, based on 1,672 cases, contains incomplete data. Allison states that the data are missing at random. As can be seen in lines 36 through 42 of appendix 2, data are available only on variables V1 and V3 (i.e., no data exist for V2 and V4). The covariance matrix and means shown in the input file is of the same dimension as in the complete data sample, that is, with four variables, to keep the notation V1 and V3 for the available data (rather than V1 and V2 for two variables, which is what EQS otherwise would assume for two input variables). The entries corresponding to V2 and V4 are completely arbitrary and have no meaning; by the model setup, EQS will not even read these entries, and only the data corresponding to V1 and V3 will be read in by the program and analyzed. The model for these variables is given in lines 28 through 35 of appendix 2. Because only V1 and V3 have data, equations for these variables only are provided. Variances and covariances are specified for factors and errors given in these equations. The final critical part of the setup lies in the cross-group constraints, which specify that every free parameter in the incomplete data sample is to be estimated at the same value as in the complete data sample.

The model was estimated with EQS, yielding χ^2_6 7.69, an acceptable model with probability p=.26. Note that there were 19 sample covariances and means to be analyzed, 20 free parameters in the model setup, and 7 cross-group constraints, yielding 19-20+7=6 degrees of freedom. The final parameter estimates, not shown here, make optimal use of all available data. In addition,

the usual output, such as a standard error for each estimate, is available for further analysis. In this example, the Lagrange Multiplier test (specified in input line 51) indicated that the constraint of equal variances of E1 across samples may not be needed (χ^2_1=4.08, p=.04). The equality of residual variances across these samples may not be important, so the model was reestimated after removing line 49 of appendix 2. The resulting model yielded an excellent fit (χ^2_5=3.21, p=.67), with a comparative fit index (Bentler, in press) of 1.00.

To conclude this section, if there are several experimental and control groups and the number of missing data patterns in any single condition is very large, the total number of groups that must be analyzed is the number of missing data patterns across all groups. Thus, the structural modeling approach becomes impractical. A major reason is that multiple-group models are hard to estimate and test, certainly harder than standard one-group structural models. Then, as in the usual approach, it may be necessary to discard those patterns of missing data that only a few subjects exhibit, to bring the problem down to a manageable size. Minimal bias will result if this loss of data does not include much selectivity bias.

INDICATORS OF PROGRAM PARTICIPATION

Prevention interventions sometimes contain a large number of specific program elements, for example, information, education, social skills training, assertiveness training, cognitive-behavioral skills training, decisionmaking training, dealing with emotion, modeling, as well as nonspecific elements such as attention from a research team. Furthermore, for each of these program components, an individual participant may be exposed to only a few elements on rare occasions and all elements on many occasions. If an experimental design breaks down and/or a quasi-experimental design is undertaken *a priori,* it may be desirable to model the consequences of the intervention in terms of the strength of the program as delivered to the individual subjects. The experimental group may be receiving many program elements associated with the controls, while the controls may be receiving many of the elements intended for the experimentals. When creating a model of the program as delivered, in essence, one would attempt to specify the mediational processes that might be at work in hindering or helping program effectiveness (Bentler and Woodward 1978; Judd and Kenny 1981).

One approach to this problem would be to use one or more latent variables to indicate exposure to program elements, considering these elements in their most minute, though identifiable, form. Suppose that 10 program elements can be identified and each subject scored in terms of exposure to each element. Several of these elements may be related to aspects of social skills training,

while others may be related to resistance of peer pressures. One might hypothesize two such exposure factors, using the observed exposure scores as indicators of the factors. On the outcome side, one might similarly create latent variables to provide more error-free indicators of program success. In this case, a complete latent variable model can be set up, with the main interest in the effects of the degrees of exposure factors on the outcome factors. The role of the latent variables is to eliminate bias due to measurement errors, which are bound to be substantial in such a situation. Stated differently, the program impact would be quantified at the latent variable level, thus having the potential to identify subtle effects that are too gross to be noticed at the level of measured variables. This methodology can be used along with other control variables in a larger structural model to help minimize errors of inference due to experimental contamination.

CONCLUSION

Structural equation modeling provides a useful approach for analyzing data from experiments that have been degraded, for analyzing nonexperimental data under hypotheses that permit control of possible confounding sources of variance in the outcomes, for separating true from observed effects when variables are measured with error and multiple indicators of latent factors are available, for efficientiy estimating parameters or testing models when data are missing at random, and for evaluating consequences of program participation when various specific program elements may be differentially reaching the intervention target population. In general, applications of modeling in intervention research require a thoughtful analysis of all the processes, intended and unintended, that may be operating to produce particular outcomes. When the analysis is thorough and the statistical assumptions are met, structural modeling can provide new insights on the intervention process, either by confirming hypothesized effects or by pointing to unexpected, but plausible, effects.

REFERENCES

Allison, P.D. Estimation of linear models with incomplete data. In: Clogg, C., ed. *Sociological Methodology 1987.* San Francisco: Jossey-Bass, 1987. pp. 71-103.

Bentler, P.M. *EQS Structural Equations Program Manual.* Los Angeles: BMDP Statistical Software, 1989.

Bentler, P.M. Comparative fit indexes in structural models. *Psychol Bull,* in press.

Bentler, P.M., and Woodward, J.A. A Head Start reevaluation: Positive effects are not yet demonstrable. *Eval Q* 2:493-510, 1978.

Bielby, W.T.; Hauser, R.M.; and Featherman, D.L. Response errors of black and nonblack males in models of the intergenerational transmission of socioeconomic status. *Am J Sociol* 82:1242-1288, 1977.
Cochran, W.G. Errors of measurement in statistics. *Technometrics* 10:637-666, 1968.
Hsu, L.M. Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. *J Consult Clin Psychol* 57:131-137, 1989.
Judd, C.M., and Kenny, D.A. Process analysis: Estimating mediation in treatment evaluations. *Eval Rev* 5:602-619, 1981.
Lord, F.M. Large sample covariance analysis when the control variable is fallible. *J Am Stat Assoc* 55:307-321, 1960.
Muthén, B.; Kaplan, D.; and Hollis, M. On structural equation modeling with data that are not missing completely at random. *Psychometrika* 52:431-462, 1987.
Rubin, D.B. Inference and missing data. *Biometrika* 63:581-592, 1976.
Sörbom, D. An alternative to the methodology for analysis of covariance. *Psychometrika* 43:381-396, 1978.
Sörbom, D. Structural equation models with structured means. In: Jöreskog, K.G., and Wold, H., eds. *Systems Under Indirect Observation: Causality, Structure, Prediction.* Vol. I. Amsterdam: North-Holland, 1982. pp. 183-195.
Weng, L.-J., and Bentler, P.M. Linear structural equation modeling with dependent observations. *Proc Soc Stat Sect Am Stat Assoc,* 1987. pp. 498-500.

ACKNOWLEDGMENT

This research was supported in part by National Institute on Drug Abuse grants DA-00017 and DA-01070.

AUTHOR

Peter M. Bentler, Ph.D.
Professor
University of California-Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024-1563

APPENDIX 1

EQS Setup for Experimental/Control Example

```
 1  /TITLE
 2  OLSSON'S DATA (SÖRBOM 1978), CONTROL GROUP
 3  /SPECIFICATION
 4  VARIABLES = 4; CASES = 105; ANALYSIS = MOMENT; GROUPS = 2;
 5  /LABELS
 6  V1 = SYNONYM1; V2 = OPPOSIT1; V3 = SYNONYM2; V4 = OPPOSIT2;
 7  F1 = ABILITY1; F2 = ABILITY2;
 8  /EQUATIONS
 9  V1 =  20*V999 +    F1  + E1;
10  V2 =  20*V999 + .9*F1  + E2;
11  V3 =  20*V999 +    F2  + E3;
12  V4 =  20*V999 + .9*F2  + E4;
13  F1 =   0V999 +    D1;
14  F2 =   0V999 + .9*F1  + D2;
15  /VARIANCES
16  D1 = 30*; D2 = 5*;
17  E1 TO E4 = 10*;
18  /COVARIANCES
19  E2,E4 = 5*;
20  /MATRIX
21  37.626
22  24.933  34.680
23  26.639  24.236  32.013
24  23.649  27.760  23.565  33.443
25  /MEANS
26  18.381  20.229  20.400  21.343
27  /END
28  /TITLE
29  OLSSON'S DATA (SÖRBOM 1978), EXPERIMENTAL GROUP
30  /SPECIFICATION
31  VARIABLES = 4; CASES = 108; ANALYSIS = MOMENT;
32  /LABELS
33  V1 = SYNONYM1; V2 = OPPOSIT1; V3 = SYNONYM2; V4 = OPPOSIT2;
34  F1 = ABILITY1; F2 = ABILITY2;
35  /EQUATIONS
36  V1 =  20*V999 +    F1  + E1;
37  V2 =  20*V999 + .9*F1  + E2;
38  V3 =  20*V999 +    F2  + E3;
39  V4 =  20*V999 + .9*F2  + E4;
```

```
F1 = 2*V999 +    D1;
F2 = 2*V999 + .9*F1  + D2;
/VARIANCES
D1 = 30*; D2 = 5*;
E1 TO E4 = 10*;
/COVARIANCES
E2,E4 = 5*;
/MATRIX
50.084
42.373  49.872
40.760  36.094  51.237
37.343  40.396  39.890  53.641
/MEANS
20.556  21.241  25.667  25.870
/CONSTRAINTS
(1,V1,V999) = (2,V1,V999);
(1,V2,V999) = (2,V2,V999);
(1,V3,V999) = (2,V3,V999);
(1,V4,V999) = (2,V4,V999);
(1,V2,F1)   = (2,V2,F1);
(1,V4,F2)   = (2,V4,F2);
(1,F2,F1)   = (2,F2,F1);
/LMTEST
/END
```

APPENDIX 2

EQS Setup for Missing Data Example

```
 1  /TITLE
 2  INCOMPLETE DATA FACTOR MODEL (ALLISON 1987)
 3  COMPLETE DATA SUBSAMPLE
 4  /SPECIFICATIONS
 5  VAR = 4; CASES = 348; ANAL = MOM; GROUPS = 2;
 6  /EQUATIONS
 7  V1 = 17*V999 +   F1 + E1;
 8  V2 = 17*V999 + 1*F1 + E2;
 9  V3 =  7*V999 +   F2 + E3;
10  V4 =  7*V999 + 1*F2 + E4;
11  /VARIANCES
12  F1 = 117*; F2 = 14*;
13  E1 = 94*; E2 = 47*; E3 = 2*; E4 = 1*;
14  /COVARIANCES
```

```
15  F1,F2 = 25*;
16  /MATRIX
17  180.90
18  126.77  217.56
19  23.96  30.20 16.24
20  22.86  30.47 14.36  15.13
21  /MEANS
22  16.62  17.39   6.65    6.75
23  /END
24  /TITLE
25  INCOMPLETE DATA SUBSAMPLE
26  /SPECIFICATIONS
27  VAR = 4; CASES = 1672; ANAL = MOM;
28  /EQUATIONS
29  V1 = 17*V999 + F1 + E1;
30  V3 =  7*V999 + F2 + E3;
31  /VARIANCES
32  F1 = 117*; F2 = 14*;
33  E1 =  94*; E3 =  2*;
34  /COVARIANCES
35  F1,F2 = 25*;
36  /MATRIX
37  217.27
38  0     1
39  25.57   0   16.16
40  0     0    0     1
41  /MEANS
42  16.98  0     6.83  0
43  /CONSTRAINTS
44  (1,V1,V999) = (2,V1,V999);
45  (1,V3,V999) = (2,V3,V999);
46  (1,F1,F1)    = (2,F1,F1);
47  (1,F2,F2)    = (2,F2,F2);
48  (1,F1,F2)    = (2,F1,F2);
49  (1,E1,E1)    = (2,E1,E1);
50  (1,E3,E3)    = (2,E3,E3);
51  /LMTEST
52  /PRINT
53  COVARIANCE = YES;
54  /END
```

Outcome Measurement Issues in Drug Abuse Prevention Studies

James H. Dwyer and David P. MacKinnon

INTRODUCTION

This chapter focuses on specific statistical and methodological issues that arise in the measurement of drug abuse and mediators of drug abuse in prevention studies. The primary focus is on experimental and quasi-experimental studies in which the intervention impact is of primary interest. However, many of the issues raised are also relevant to population studies.

EXTRASCIENTIFIC ISSUES

When choosing a measurement strategy, it is important for researchers planning prevention studies to consider several extrascientific factors, including costs, ethics, and confidentiality. Costs encompass respondent time and personnel for administration and processing (laboratory analysis, keypunching, etc.). Ethical issues incorporate rights of privacy, invasiveness of the procedure, and associated health risks. The issue of confidentiality involves establishing procedures of sufficient security to provide adequate assurance that information cannot be linked with individual subjects.

The importance of these extrascientific issues may preclude the use of some types of measures when the specific research context is considered. The unfortunate task confronting the prevention researcher is then to determine whether a measure with more error or potential bias is adequate to achieve the desired study goals. Some measure of uncertainty always characterizes empirical studies. A major scientific goal is minimization of that uncertainty, but this minimization must be achieved within the constraints placed on study design by these extrascientific factors. No scientific rules for these judgments are available because they involve ethical and political issues as well as issues of measurement error and bias.

METHODOLOGIC ISSUES

The remaining sections of this chapter concern methodologic issues of outcome definition, measurement error, and measurement bias. The abstract definition of an outcome in the area of drug abuse prevention is almost unconstrained. That is, a researcher is free to choose between physical, biological, psychological, or social entities as outcomes. Measures of these entities may involve real numbers, integers, ordered categories, unordered categories, or other sets. The measurement of an entity, however, must involve an objective operation that maps observable events into a subset of the real numbers.

For example, suppose that the dimension ξ is defined such that event F_{it} (individual=i, time=t) is mapped into the real number $\xi(F_{it})$. Now suppose that a measurement procedure X maps F_{it} into the real number x_{it}:

$$X(F_{it})=x_{it}$$

If X involves *measurement error,* then

$$X(F_{it})=\xi(F_{it})+\delta_{it}$$

If the range of X is continuous and X is *biased,* then

$$E(\delta_{it})=\lim_{n\to\infty} \Sigma_{j=1}^{n}(\delta_{ijt})/n \neq 0$$

for some t>0 or some i, where j indexes independent measurements. If the bias, $E(\delta_{it})$, is the same for all i and t, then the measure is unbiased in the relative sense and is appropriate for comparisons between subgroups of the population. The primary goal of measurement, then, is to find a measurement operation for a given entity (latent variable) that minimizes both the amount of error (indexed by the variance of δ_{ijt} over j, $\sigma\delta\delta$) and the bias. When sufficient minimization of error variance is not achievable, then multiple unbiased measures can be used to adjust regression parameters for the measurement error.

If the range of X is not continous, then an alternative formulation is more convenient. Suppose the range of X is ordered categorical such that $X(F_{it})=0, 1$

then $\Phi[X]=\xi(F_{it})+\delta_{it}$). Where Φ is some transformation (e.g., probit or logit) of the probability that X (F_{it})=x_{it}, the mapping ξ is then the "true" probit or logit of the probability of event F_{it}.

Prevalence vs. Incidence

Prevention researchers often have conceptualized their interventions as changing the incidence of drug use, that is, the transition from nonuser to user (usually in adolescence). The notion of incidence is taken from epidemiology in which the first transition from nondiseased to diseased status, is the event of interest.

Incidence

The standard statistical model for estimating the dependence of incidence rates on experimental or observed variables is the proportional hazards model (Cox 1972), which is a continuous time form of the logistic model (Breslow, in press). The hazard rate is best understood by beginning with the differential equation describing the survival curve, y(t). The survival curve is equal to the population size at time zero: y(0)=N. As time passes, members of the population make the transition. The slope of the survival curve at t, dy(t)/dt, is then a measure of how rapidly the curve is declining (i.e., mortality events per unit time). The hazard rate, h(t), then is defined as the ratio of dy(t)/dt and the number of survivors to time t, y(t) (multiplied by minus one, so that the hazard rate is positive):

$$h(t)=-[1/y(t)]dy/dt$$

$$\text{or } h(t)=\lim_{\Delta t\to 0}\{y(t)-y(t+\Delta t)\}/\Delta t y(t)=\lim_{\Delta t\to 0}\frac{\text{prob. (mortality before } t+\Delta t/\text{survival to } t)}{\Delta_t}$$

The hazard rate is then minus one times the slope of the survival curve divided by its height, the instantaneous rate of change in the probability of failure per unit time, given survival to t. The rationale for this formulation is that the "hazard" inherent in the rate of decline in the survival curve (-dy/dt) increases as the height of the curve [y(t)] decreases.

The integrated form of the hazard model depends on the form of y(t).
If $y(t)=Ne^{-\beta t}$, then h(t) is constant and equal to β because $dy/dt=-\beta Ne^{-\beta t}=-\beta y(t)$. In this instance the probability of mortality between time t_1 and t_2 ($t_2>t_1$), conditional on survival to t_1, is then $[y(t_1)-y(t_2)]/y(t_1)$ or $[1-e^{-\beta(t_2-t_1)}]$, which is zero when $t_1=t_2$ and approaches one as (t_2-t_1) increases.

The advantage of the continuous time hazard formulation over discrete-time methods is that differences between groups in h(t) are independent of the length of followup used to estimate h(t) for each group. This is clear in the case where h(t) is presumed constant over time in each group. In Cox's "proportional" formulation of the hazard model, the form of y(t) is not specified. However, it is assumed that the shape of y(t) for the groups being compared [e.g., $y_0(t)$ and $y_1(t)$] is such that $h_1(t)/h_0(t)$ is constant for all t. This line of reasoning can be extended to the case in which $h_x(t)/h_{x=0}(t)$ is constant over t but varies with values of the variable x. The form of the covariance between the ratio of hazards and x often is assumed to be exponential:

$$h_x(t)/h_{x=0}(t)=e^{\beta x}$$

where the intercept is necessarily zero (because the ratio of hazards is 1 when x=0).

Application of the proportional hazards model to drug abuse prevention studies is appropriate when a transition from a universal category (nonuser at a young age) to an absorbing category (user) is of interest. However, unlike mortality, drug use is either a set of categories that can be entered and left numerous times during the course of a lifetime or a continuum along which individuals can move up or down through time.

An alternative epidemiological concept to incidence that generally will be of greater utility in drug abuse prevention studies is that of *prevalence* or *prevalence rate.* Given a geographically defined population or a tracked panel of size N(t) at time t, the prevalence of the characteristic Y (=0 or 1) at time t, y(t), is the number of persons with Y=1 at t. Therefore, the prevalence rate is the proportion of the population, y(t)/N(t), with the characteristic Y at time t.

In the context of a drug abuse prevention study, the prevalence rate, y(t), of drug use (Y=1) at time t is the population or panel characteristic of primary concern. The public health goal of prevention studies is to estimate any reduction in y(t) in an intervention condition relative to a control condition. Statistical models for this purpose are reviewed below. Researchers concerned with testing for differences in program effectiveness between baseline users and nonusers should use these models to test for an interaction between baseline status and program effect. Measurement of use among baseline nonusers often is described in terms of incidence; however, the first instance of drug use seldom defines the outcome of interest.

Continuous Dimensions, Categories, Ordered Categories, Stages, and Indices

There are numerous potential outcomes in the evaluations of drug abuse prevention interventions. The continuous dimension "level of exposure" to a substance is probably the most straightforward and is of primary interest in epidemiologic studies concerned with the health consequences of drug use. However, studies concerned with the social psychology of becoming a drug abuser may measure dimensions, categories, or stages of drug use that are defined primarily by personal and social perceptions.

Multiple Drug Outcomes

When defining drug use outcomes involving use of several drugs, it is important to distinguish between alternative models:

- *Separate drug model.* Each drug is treated as a separate variable with potentially unique determinants and consequences.

- *Polydrug model.* It is supposed that a general drug use dimension is reflected in the use of various drugs. Determinants and consequences of change in drug use levels occur because of change in the polydrug dimension, and an increase in this dimension is indicated by increased use of all drugs included.

- *Index model.* In contrast to the polydrug model, the index model does not necessarily predict an across-persons association between the level of use for different substances. Rather, level of use on the index model is a sum of use across drugs. Therefore, different drugs are substitutable for one another.

The polydrug dimension is specified statistically in terms of a factor model in which use levels for drugs are functions of the polydrug dimension. A drug use index is computed by summing drug use items. The index model treats different drugs as substitutable in an additive sense. The polydrug model treats different drugs as equally likely manifestations of the polydrug latent variable.

The importance of explicit specification of the measurement model and a clear understanding of its meaning show that the same data can yield very different conclusions depending on the model assumed. Furthermore, the data collected in a study may be insufficient to distinguish among the alternative models.

The potential for these conflicting interpretations is demonstrated by the following example. Suppose that continuous measures of three drugs (Y_1, Y_2,

Y_3), a measured exogenous variable (X), and an unmeasured variable ξ covary because of the following causal model:

$$
\begin{aligned}
Y_3 &= \gamma_1 \xi + \zeta_1 & \gamma_1 &= .5 \\
Y_2 &= \gamma_1 \xi + \zeta_2 & & \\
Y_1 &= \beta X + \gamma_1 \xi + \zeta_3 & \beta &= .3 \\
X &= \gamma_2 \xi + \zeta_4 & \gamma_2 &= .4
\end{aligned}
$$

where all variables have mean zero and unit variance (except the disturbances ζ) and cov (ξ_4,X)=0. This causal model states that X influences use of Y_1, but X has no impact on the other two drugs. Use of the three drugs also is correlated because of an unmeasured background variable ξ (e.g., geographic region, socioeconomic status, etc.). This is a plausible model for many observed variables X, an experimental manipulation of X (γ_2=0) or for a quasi-experimental study ($\gamma_2 \neq 0$).

Assuming that the disturbances are uncorrelated, this causal model implies the following correlation matrix among the variables:

	Y_1	Y_2	Y_3	X
Y_1	1.00			
Y_2	0.31	1.00		
Y_3	0.31	0.25	1.00	
X	0.50	0.20	0.20	1.00

Now assume the following polydrug latent variable model:

$$
\begin{aligned}
Y_1 &= \tau_1 + \eta_1 + \varepsilon_2 \\
Y_2 &= \tau_2 + \lambda_1 \eta_1 + \varepsilon_2 \\
Y_3 &= \tau_3 + \lambda_2 \eta_1 + \eta_3 \\
\eta_1 &= \beta_1 X + \zeta_1 \\
\eta_2 &= \beta_2 X + \zeta_2
\end{aligned}
$$

where η_1 is the polydrug latent variable; η_2 is a combination of measurement error in Y_3 and drug use variance that is not explained by the polydrug latent variable; and the measurement errors (ε) and the disturbances (ζ) are

presumed to be random. If we first constrain β_2=0 (the simple polydrug model), the value 0.492 is obtained for the path β_1. The conclusion then would be that X has an impact on all drugs via the polydrug latent variable. If we then allow β_2 to be a free parameter, we find that β_1=0.40 and β_2=0.378. This latter finding would lead the researcher to conclude that X had an impact on both the polydrug factor and on the specific variance in Y_1. The model that generated the data, however, involved an impact of X only on Y_1.

This exercise emphasizes the care that must be applied in the interpretation of polydrug models. In view of these difficulties and that randomized interventions may affect various drugs differently, the use of the polydrug latent variable is somewhat problematic.

Quantity and Frequency of Use. Although the polydrug and index models simplify the data by combining information into unitary entities, there are reasons to consider moving in the opposite direction toward even greater specificity. In the case of alcohol use, for example, interventions may change self-reports of the amount consumed when drinking, even though frequency of alcohol use is unaffected. Thus, it may be of importance to the progress of prevention research to distinguish between the prevalence of heavy use (when using) and the prevalence of frequent use (at any level).

Multiple Indicator Models of Measurement Error

The most developed type of multiple indicator model of measurement error is based on the latent variable or factor model (Dwyer 1983). These models deal with the "errors-in-variables" problem in regression models. Even if measurement error (ME) in a predictor variable has an expected value of zero and is uncorrelated with other variables in a system, such ME gives rise to bias in estimates of regression slopes. Such ME in a dependent variable does not, however, bias estimates of regression slopes when the dependent variable is continuous (unless variables are standardized). When the dependent variable is ordered categorical, then error will bias slopes in logistic or probit regression models. Thus, it can be of considerable importance to include multiple indicators in a study so that bias in estimates of regression coefficients can be removed.

A second reason for incorporating a multiple indicator measurement model in drug abuse prevention studies is that an explicit test of self-report bias can be performed by incorporating self-report and biological measures in a single model. The importance of evidence to counter the "report bias" alternative explanation of intervention effects is a judgment that must be made by each investigator.

Types of Measurement Models. Multiple indicator measurement models (MIMMs) have been developed for an array of variable types (Arminger and Kusters, in press). Software is widely available to estimate measurement models for the following variable types:

- Continuous, normally distributed variables
- Continuous, arbitrarily distributed variables
- Ordered categorical variables

For the continuous, normal case, Jöreskog and Sörbom's (1988) LISREL model and software are well known. For the continuous, nonnormal case, two options are available. The first is the pseudomaximum likelihood (PML) estimation procedure developed in econometrics (Arminger and Schoenberg 1989); this approach uses the difference between the square of the first partial derivatives and the second partial derivatives of the log likelihood function to detect nonnormality and adjust standard errors. The second approach is the asymptomatically distribution-free (ADF) estimation procedure (Browne 1984), where nonnormality is detected via univariate kurtoses. The ADF approach has been implemented in LISREL 7 and EQS (Bentler 1986). The PML approach is available in the program LINCS (Schoenberg 1987), which is written in the GAUSS language (Edlefson and Jones 1986). The PML approach has some practical advantages over ADF, but the use of LINCS is currently cumbersome relative to LISREL and EQS.

Maximum likelihood estimation of measurement models in the case of ordered categorical variables was developed by Muthén (1979) and has been implemented in the program LISCOMP (Muthén 1987). Version 7 of LISREL also includes a preprocessor that computes input for a multivariate probit model. The estimation procedure employed is ADF. The point estimates from the LISREL procedure are close to those obtained from LISCOMP in large samples if the LISREL model is specified as a probit (the disturbance variance is fixed at one). However, the standard errors from LISREL may be biased toward zero.

Biochemical Indicators

For some researchers, demonstration of program effects on self-report is inadequate because of potential confounding by recall bias. For example, participants in a program may be less likely to report drug use. In the case of cigarette use, several biochemical indicators are available, including carbon monoxide in expired air and thiocyanate and cotinine in body fluids. These measures then can be included in MIMMs to assess fit of the model or used separately as dependent variables.

Conditional vs. Unconditional Models of Change in Longitudinal Studies

Most drug abuse prevention studies involve baseline and one or more followup measurements after some units such as schools receive the prevention program. Such longitudinal designs allow the study of change in drug use over time, thereby filtering out confounders of program effects that do not change over time. Longitudinal studies, however, do not remove time-varying factors that may confound program effects. Additional followup measurements provide the opportunity to examine the temporal sequence of relevant variables such as whether changes in program-mediating variables precede changes in drug use or vice versa (Dwyer, in press).

Prevention studies often involve barriers to randomization of units to experimental conditions. Administrators may want to implement the program immediately rather than allow some units to serve as comparison for changes in other schools. However, individual administrators may demand program or control status. In the absence of randomization, the magnitude of intervention effects should be evaluated under alternative assumptions about the causal process that may generate baseline nonequivalencies. For example, what is the magnitude of the program effect with allowance for regression to the mean (the conditional model)? Alternatively, if nonequivalencies are a stable characteristic of the units under study, then the magnitude of the program effect under the assumption of maintenance of pretest differences should be evaluated (the unconditional model).

The conditional and unconditional models are summarized in the following equation:

$$y_{i1}=\alpha_0+\beta_1 y_{i0}+\beta_2 x_i+\zeta_i$$

where x is a dichotomous dummy variable indicating experimental condition (treatment or control); α_0 is the intercept. The constraints $E(\zeta)=\sigma_{x\zeta}=\sigma_{y_0\zeta}=0$ specify the conditional version of the model; and the constraints $\beta_1=1$, $E(\zeta)=\sigma_{x\zeta}=0$ specify the unconditional version. The value (β_1-1) reflects the speed with which the dependent variable regresses to an equilibrium level. When $\beta_1=1$, the dependent variable does not regress to mean levels and is equivalent to regression on the difference $(y_{i1}-y_{i0})$ dependent variable (the unconditional model). The unconditional model also may be specified as a multiple equation regression for the case of more than two followup measurements.

When program and control groups are equivalent at baseline, the conditional and unconditional models yield identical estimates of program effects. That is, the conditional and unconditional models differ in their definition of β_2 only when $\sigma_{y_0x} \neq 0$:

$$\beta_2 = (\mu_{y_1|x=1} - \mu_{y_1|x=0}) - \beta_1(\mu_{y_0|x=1} - \mu_{y_0|x=0})$$

where $\beta_1 = 1$ in the unconditional case and β_1 is the within-group regression of y_1 on y_0 in the conditional case. Thus, in the case of randomization of a large number of units to conditions, differences between program and control groups are assumed to be the result of random sampling error. Applications of conditional and unconditional models in drug prevention research are described by Dwyer and associates (1989).

An important, underutilized design is to include two or more baseline measurements to provide information on whether baseline equivalencies either maintain or regress to mean levels. In this way, the time course that group differences would have taken in the absence of an intervention is based on the preprogram measures (Dwyer 1984).

SUMMARY

Like other areas of applied research, there are extrascientific issues such as economics, ethics, and confidentiality in prevention research. After acknowledging extrascientific issues in the design and implementation of research programs, several improvements were suggested to evaluate drug abuse prevention efforts. Prevalence rather than incidence of drug use is a useful dependent variable in drug prevention research because of an individual's transition in and out of drug use. When a prevention program changes one drug, but analysis is conducted on an index of several drugs, program effects may be misleading. Thus, it is suggested that prevention program evaluation may be improved if effects on different drugs and different levels of use are assessed.

Multiple measures of constructs in drug prevention will improve the reliability of these constructs. In this regard, biological measures add substantially to the reliability and interpretability of prevention program effects. Estimation of prevention program effects incorporating the categorical nature of drug use may increase understanding and concretize prevention program effects. New software for estimating models with multiple categorical and continuous

measures under different distributional assumptions should increase the application of these analysis techniques.

Finally, prevention researchers often are confronted with nonrandom assignment of units to conditions. In this situation and when the success of randomized assignment is questionable, it is important to evaluate program effects under different assumptions to examine what would happen in program and control groups in the absence of a prevention effect. The conditional and unconditional models assume that program and control groups either regress to mean levels or maintain pretest differences, respectively.

REFERENCES

Arminger, G., and Küsters, U. Toward a general model for longitudinal change in categorical and continuous variables observed with error. In: Dwyer, J.H.; Feinleib, M.; Lippert, P.; and Hoffmeister, H., eds. *Statistical Models for Longitudinal Studies of Health.* New York: Oxford University Press, in press.

Arminger, G., and Schoenberg, R.J. Pseudo maximum likelihood estimation and a test for misspecification in mean and covariance structure models. *Psychometrika* 24:409-425, 1989.

Bentler, P.M. *Theory and Implementation of EOS: A Structural Equations Program.* Los Angeles: BMDP Statistical Hardware, 1986.

Breslow, N.E. Logistic models for longitudinal studies. In: Dwyer, J.H.; Feinleib, M.; Lippert, P.; and Hoffmeister, H., eds. *Statistical Models for Longitudinal Studies of Health.* New York: Oxford University Press, in press.

Browne, M.W. Asymptomatically distribution-free methods for the analysis of covariance structures. *Br J Math Stat Psychol* 37:62-83, 1984.

Cox, D.R. Regression models and life tables. *J R Stat Soc* 34B:269-276, 1972.

Dwyer, J.H. *Statistical Models for the Social and Behavioral Sciences.* New York: Oxford University Press, 1983.

Dwyer, J.H. The excluded variable problem in nonrandomized control group designs. *Eval Rev* 8:559-572, 1984.

Dwyer, J.H. Overview of statistical models for longitudinal data. In: Dwyer, J.H.; Lippert, P.; and Hoffmeister, H., eds. *Statistical Models for Longitudinal Studies of Health.* New York: Oxford University Press, in press.

Dwyer, J.H.; MacKinnon, D.P.; Pentz, M.A.; Flay, B.R.; Hansen, W.B.; Wang, E.; and Johnson, C.A. Estimating intervention effects in longitudinal studies. *Am J Epidemiol* 130(4):781-795, 1989.

Edlefson, L.E., and Jones, S.D. *GAUSS Programming Language Manual: Version 1.49B.* Kent, WA: Aptech Systems, Inc., 1986.

Evans, R.; Hansen, W.B.; and Mittlemark, M. Increasing the validity of self-report of smoking behavior in children. *J Appl Psychol* 69:521-523, 1977.

Jöreskog, K., and Sörbom, D. *LISREL VII: A Guide to the Program and Applications.* Chicago: SPSS, Inc., 1988.
Muthén, B. A structural probit model with latent variables. *J Am Stat Assoc* 74:807-811, 1979.
Muthén, B. *LISCOMP: Analysis of Linear Structural Equation Models Using a Comprehensive Measurement Model.* Mooresville, IN: Scientific Software, 1987.
Schoenberg, R. *LINCS: Linear Covariance Structure Analysis Users Guide.* Kensington, MD: RJS Software, 1987.

ACKNOWLEDGMENT

This work was supported by Public Health Service grant DA-03976 from the National Institute on Drug Abuse.

AUTHORS

James H. Dwyer, Ph.D.
Associate Professor of Preventive Medicine
Institute for Health Promotion and Disease Prevention Research
Suite 641
1000 South Fremont Avenue
Alhambra, CA 91803-1358

David P. MacKinnon, Ph.D.
Associate Professor
Department of Psychology
Arizona State University
Tempe, AZ 85287-1104

Assessing Effectiveness of Drug Abuse Prevention: Implementation Issues Relevant to Long-Term Effects and Replication

J. David Hawkins, Robert Abbott, Richard F. Catalano, and Mary R. Gillmore

INTRODUCTION

Ultimately, drug abuse prevention interventions must be assessed for their success in preventing drug abuse and its associated costs. The term "drug abuse" has been used to describe at least six empirically and conceptually distinct types of drug-related behavior, ranging from a single episode of substance use to repetitive pathological use over an extended period (Hawkins et al. 1985; Hawkins and Catalano 1989). Although the intervention effects on several indicators of these drug use outcomes have been investigated, little is known about the effects of drug abuse prevention interventions on patterns of pathological drug use that persist for longer than a month and cause social- or occupational-impaired functioning in the family, at school, or in a work setting (American Psychiatric Association 1980). Virtually no prevention evaluations have followed participants for an extended period to assess long-term effects; yet, it is only through long-term replication studies that the effects of drug abuse prevention interventions will be revealed.

The most complete evaluation data currently come from studies of classroom-based interventions focused on developing resistance to social influences to use drugs, although at least one recent study has broadened the intervention setting to include the use of media, community, and parent involvement in the prevention strategy (Pentz et al. 1989). These prevention interventions focus primarily on reducing two identified drug abuse risk factors: social influences to use drugs and social norms favorable to drug abuse. Some studies have shown significant short-term reductions in the onset and prevalence of cigarette smoking following resistance training in comparison with controls (Botvin 1986). A few studies have shown reductions in the prevalence (within 18 months of

intervention) of alcohol and marijuana use among those exposed to a combination of extensive instruction in social influence recognition, resistance skills, and content-seeking to promote negative attitudes toward the use of all drugs when compared with nonexposed subjects (Botvin 1987; Hansen et al. 1988; Pentz et al. 1989).

Unfortunately, in the few studies of school-based social influence resistance programs that have followed subjects 2 years or more beyond the intervention, initial posttest differences in the prevalence of cigarette, alcohol, and marijuana use that initially favored experimental groups disappeared (Hansen et al. 1988; Botvin 1987). Available evaluation studies provide little understanding about the deterioration in effects that were initially observed.

Two major hypotheses are tenable. First, it is possible that drug use patterns 2 or more years following preventive intervention are influenced by other factors that overwhelmed the intervention over time. In this hypothesis, the lack of longer term effects is attributed to a weak intervention, which, by itself, might not be expected to have sizable or long-term effects. This hypothesis suggests that prevention research should explore the effects of interventions targeted on other identified substance abuse risk factors to supplement social influence resistance strategies (Hawkins et al. 1989). Second, the apparent absence of longer term effects might be taken as evidence of a potentially strong intervention for which the "dose" administered was inadequate. This hypothesis suggests that, in these prevention programs, social influence resistance skills and antidrug attitudes were not taught well enough to ensure their maintenance over a longer timeframe. This second hypothesis suggests that researchers increase the implementation level of the resistance skills intervention, with periodic booster sessions and other activities.

In summary, there is not enough evidence to determine whether drug abuse prevention interventions actually prevent drug abuse when the American Psychiatric Association definition is used. In addition, where followup evaluations of prevention interventions have been conducted, there is insufficient evidence to determine whether the effects of evaluated drug abuse prevention strategies have deteriorated over time because these strategies were based on inadequately specified theories or because the interventions were not adequately implemented to produce the desired longer term effects on drug use behaviors. For future policy and research, we need to know why effects occur in prevention evaluation studies and whether the interventions are effective.

This chapter outlines a strategy for assessing the long-term effects of drug abuse prevention interventions in replicable studies.

THEORY CONSTRUCTION

A first step in enhancing the possibility of replication and assessment of the long-term effects of a preventive intervention is specifying the intervention theory. This theory must define the theoretical basis for the intervention and the target intervention audiences and suggest the short- and long-term intervention outcomes. In addition to defining these constructs, the theory must specify linkages among the intervention, target audiences, and outcomes. Defining these cor structs and describing their theoretical linkages are the first step in theory-testing, replication, and analysis of long-term followup effects.

With respect to replication, the definition of key theoretical constructs is of critical importance. To illustrate, social influence resistance interventions have been based on a drug abuse theory that includes the hypothesis that initiation and drug use during early adolescence results from inadequate skills to recognize and resist social influences to use drugs. The interventions tested in these studies include refusal skills training. In this illustration, the key theoretical construct is skills to resist drug influences, which is the construct that interventions seek to change. Studies that seek to test interventions based on this hypothesis should measure the theoretically relevant construct of resistance skills. When comparing results across studies, it is important to compare the extent of short-term effects on resistance skills to determine whether each intervention was conducted in a comparable fashion and whether the intervention effected the constructs of interest. It is not essential to show that the intervention technology was exactly the same across studies. This suggests an interest in common construct indicators that have theoretical relevance for drug abuse prevention experiments. However, the relevance of congeneric indicators used to measure the underlying construct at different developmental points also must be recognized in long-term prevention followup studies.

The theoretical basis for the intervention also must consider the relationship between the intervention and other relevant "causal" factors that might influence the outcome. Sometimes these other factors may operate relatively independently of the intervention, and in other cases, such factors may interact with the intervention to influence outcome. This is especially true for long-term outcomes when other factors and events intervene between the preventive intervention and the measurement of long-term outcome.

To illustrate, the adolescent social development model of drug use specifies desirable risk-focused intervention points along the developmental continuum from early childhood through midadolescence. The model hypothesizes, for example, that interventions to improve teachers' instructional skills should

increase children's involvement in school and school rewards, thereby strengthening educational commitment, increasing bonding to school, and reducing the likelihood of initiating drug use in the elementary grades (Catalano and Hawkins 1986). However, the model also specifies the importance of family management, peer involvement, and problemsolving skills in predicting the incidence, extent, and frequency of drug use during the middle school period.

We have tested interventions focused on improving elementary school teachers' instructional skills as a strategy for preventing the early initiation of drug use and have found lower rates of drug use initiation at fifth grade entry for students exposed to the teaching interventions for a semester or more in grades 1 through 4 (Hawkins et al. 1989). Nevertheless, we hypothesize that the intervention's effects on drug use by the end of middle school will be indirect, operating largely through effects on early initiation of drug use, through retention and involvement in different classes in secondary school, and in turn, through positive peer associations formed during that period. Using forms associated with structural modeling, these hypotheses suggest that there are no direct effects of instructional intervention in the elementary grades on drug use in grade 9 but direct intervention effects will surface if constructs are specified and measured.

To better understand the effects of the instructional intervention, the influence of other causal factors defined by the model must be estimated. Failure to include all theoretically relevant constructs for testing of the long-term effects of the instructional interventions will result in specification errors that confound accurate assessment of the intervention effectiveness (Costner 1971).

Prevention evaluations based on conceptual foundations that take cognizance of adolescent substance use risk factors will become, almost out of necessity, similar to etiologic studies because both these types of studies seek to identify factors and processes that increase or mitigate the likelihood of drug initiation and use. This similarity is the result of prevention evaluation studies that consider all factors that are theoretically hypothesized to contribute to the outcomes or dependent measures. These factors should be measured and modeled if the intervention's empirical effects are to be understood. Therefore, the implication is that prevention and etiologic studies should become less distinct both in the measures used and the analytic strategies applied to data.

Testing theory involves comparing the consistency of theoretical linkages between intervention and outcome with the empirical linkages present in the data. Consistency between the theory and data provides support for the theoretical linkages between intervention and outcome. The research design

and statistical analysis also must attempt to rule out competing explanations for the empirical results other than those suggested by the theoretical predictions. For example, Biglan and Ary (1985) documented possible attrition effects on the outcome of smoking prevention interventions. The need to rule out other competing interpretations is present whether the data are consistent or inconsistent with the theoretical linkages. In either case, researchers can examine and rule out a large number of alternative explanations for their results by carefully assessing the degree to which interventions were implemented with fidelity by incorporating the information on degree of implementation into the analysis of the linkages among the intervention, modeled factors, and outcomes.

Careful assessments of the degree to which an intervention was implemented and the inclusion of these assessments in the statistical analysis of the interventions efficacy will enhance the replicability and analysis of the long-term effectiveness of the intervention. Research tasks to carry out these assessments include collection, reporting, and analysis of the data on the integrity and fidelity of program implementation and the incorporation of these data into efficacy tests of the intervention.

COLLECTING DATA ON DEGREE OF IMPLEMENTATION

Many methods of collecting data on implementation have been proposed. For some studies, researchers have used records of dates and places of intervention sessions along with the names of those in attendance to judge the integrity of the intervention. Other researchers have used audio and video recordings of intervention sessions or spot-checking of intervention sessions by supervisors. Others investigating the efficacy of school-based intervention approaches have relied on teacher self-reports or teacher interviews (Hall and Loucks 1977; Shaver 1983).

Researchers also have assessed whether subjects accurately perceive intervention characteristics that should be obvious if preventive interventions differ as intended. For example, one of the instructional methods included in our prevention work with teachers is the use of cooperative learning in which students work in small classroom teams to master the subject matter (Hawkins et al. 1988). As hypothesized, when surveyed at the end of an academic year of intervention, experimental students in comparison with controls reported significantly greater agreement with the statement, "In my classes, we break up into groups which compete with each other" (Hawkins and Lam 1987). Had this difference not been observed, it would indicate that students' perceptions were inconsistent with assumptions about how the intervention was affecting students. Results indicating no intervention effect are hardly surprising in such cases (Leinhardt 1980).

A common method of collecting data on the degree of intervention implementation is to use observational methods, which include informal observation, global judgments made by observers blind to whether they are observing the treatment or control intervention, anecdotal reports, qualitative analyses based on long-term ethnographic study of the intervention, and ratings based on systematic observation.

Researchers who have assessed the integrity of treatment implementation by using ratings based on systematic observation have either adapted an available instrument or developed a new instrument. Although adapting an available instrument may be appropriate to study some interventions, this often results in using an observation instrument with inadequate construct validity. Adapting an observation instrument developed from some other theoretical perspective to assess interventions often results in the observer focusing on irrelevant aspects of the intervention. For example, observation systems based on Flanders' interaction analysis model have been widely used to assess the integrity of an intervention even when Flanders' theoretical basis was irrelevant to the intervention being studied.

Researchers who develop new observational systems face important measurement issues. The observational system must provide for the assessment of the critical dimensions of the intervention. These critical dimensions should be drawn from the theoretical basis for the intervention. Observers must be trained to provide consistent ratings and be given sufficient observational opportunities so that they can obtain stable estimates of the degree of implementation. If interventions are complex, observers must assess the interventions on the multiple dimensions. Observations are usually costly investments that take into account the development and refinement of observational systems.

For example, to study the effects of teacher instructional skills on student outcomes, we developed a classroom observational coding system that includes codes for 11 distinct teaching behaviors that are ordered on a 1-minute time-sampled basis. This system required extensive psychometric work to develop an interactive teaching map that could reliably distinguish teaching behaviors (Kerr and Cummings 1982; Kerr et al. 1985). The investment in this observational system has paid off. Observational data from the interactive teaching map have been useful in documenting the degree of implementation of experimental instructional methods in experimental and control classrooms and in showing the predictive power of certain instructional practices on prosocial and antisocial behavior, including school drug use (Hawkins and Lam 1987).

Note that this observational system measures only the teacher-based component of the comprehensive prevention intervention strategy that is tested. Similar observational work needs to be carried out for families of experimental children participating in the parenting education programs. Again, the theory-based strategy for intervention design and implementation measurement implies greater measurement and data collection costs than is typical in many prevention studies.

The internal validity of the study also can be strengthened by observing intervention dimensions that are not expected to affect outcomes. In this way, alternative explanations for intervention effects can be ruled out. For example, a recent analysis of cumulative exposure effects to all interventions in the Seattle Social Development Project for at least one semester in grades 1 through 4 included indicators of theoretical constructs that were influenced by teaching and parenting interventions (such as student perceptions of rewards from schooling, student commitment to schooling, and family management practices at home, and indicators of constructs that were not expected to change as a result of the preventive interventions implemented in grades 1 through 4). These latter constructs included the perceived risks of drug use, which have been hypothesized to influence drug use behaviors (Johnston 1985). By measuring constructs expected to change as a result of the intervention and those not expected to change, the study could examine whether the intervention was accompanied by halo effects. However, this does not appear to be the case because the intervention produced no significant differences between experimental and control subjects in perceived risks of drug use, although several targeted risk factors were significantly different between groups (Hawkins et al. 1989). The inclusion of data on perceived risks of drug abuse also indicated that, although the fifth grade experimental and control subjects differed significantly on the prevalence of delinquency and drug use initiation, they did not differ significantly on perceived risks of drug use. It is possible that this construct may be salient as a risk factor for drug initiation in later childhood or early adolescence. This process is similar to the use of multiple baseline designs in single subject research.

Ideally, several methods for measuring the degree of implementation should be included. Using multiple indicators provides a more complete assessment in implementing the intervention and should be used in subsequent analyses.

REPORTING IMPLEMENTATION DATA

Researchers should provide sufficient detailed implementation data to give others a clear picture of intervention differences and to allow for replication. Researchers also should provide a synopsis of implementation data and refer to

a detailed data source that provides information about critical dimensions relevant to implementation of each intervention component. Often, a matrix showing the degree of implementation of each critical intervention dimension can be displayed (Leithwood and Montgomery 1980). Although these recommendations often run counter to the desires of journal editors, such information is necessary for other researchers attempting to replicate the intervention in another setting or with another target group. Subsequent metaanalyses of effect sizes also should incorporate analysis of the degree to which an intervention was implemented. It is likely that the average effect size is related to the degree of implementation for an efficacious intervention. Presenting data on the varying degrees of component implementation for complex interventions allows subsequent metaanalyses to investigate the efficacy of various components (or their interactions) within a complex multidimensional intervention.

Systematic assessment of intervention implementation also can help identify factors that threaten the implementation process. Such factors can include personality characteristics, interests, motivational characteristics, cognitive belief systems, and other characteristics of program implementors. Situational variables also may interact with the intervention and influence the degree of implementation. Interventions delivered in the environment outside the laboratory will not be as standardized as they are in the laboratory. Prevention interventions are often complex, may be delivered by poorly trained or unmotivated people, and can be totally disrupted by events outside the study. When examining the degree of implementation, the political context of the intervention also must be considered. Once collected and reported, data on the degree of implementation of an intervention should be incorporated into the outcome analysis associated with the intervention.

ANALYZING DATA ON THE DEGREE OF IMPLEMENTATION

Basic descriptive statistical information related to implementation helps a researcher judge the degree of operationalization of the theoretical construct. Selection of appropriate implementation indicators should be based on the theoretical underpinnings of the intervention. Selection of an aggregation unit also should be based on the theoretical linkages between intervention and outcomes. The more data available about an implementation, the clearer the interpretation. For example, assume that, when examining the implementation of two interventions, a researcher discovers that one intervention is being delivered by experienced practitioners and the other by newly trained practitioners. Knowing this information about an intervention implementation allows a researcher to construct a confounding hypothesis. Results lead to the conclusion that the theoretical constructs implemented in the interventions are

differentially efficacious, namely that the experience of the practitioner, and not the intervention approach, accounts for the difference in outcomes.

When comparing the degree of implementation for different interventions, inferential statistical analysis can help a researcher examine the likelihood that differences are due to chance. Because most statistical methods are affected by sample size, the researcher must be cautious because mean differences and significance levels may mask large implementation variability. Large sample sizes may produce small degrees of implementation inconsistency that are statistically significant. For example, even though the difference between means may be statistically significant, there may still be little clinical or practical difference between the intervention and control condition. Threats to the statistical conclusion validity of the analysis must be considered carefully.

When examining the relationship between an intervention and hypothesized outcomes, the statistical analysis must assess alternative hypotheses. For example, the intervention may be confounded with an associated variable. Nonspecific intervention effects, such as expectancy or placebo effects, may be present and may be mistaken for actual effects. The intervention may not have been reliably implemented, and some individuals may not have received the full degree of the intervention. This is especially prevalent in interventions seeking to involve parents in prevention activities (Fraser et al. 1988).

When typically applied, statistical analyses based on t-tests and analyses of variance assume that the intervention conditions are fixed and have been implemented equally for all individuals. Control for alternative explanations of the results derives from random assignment of subjects to treatments and from including either large numbers of subjects in the study or multiple measures on a few subjects.

Using multiple regression techniques relating the degree of implementation to outcomes provides one method of incorporating the degree of implementation into an analysis of the efficacy of an intervention. With these techniques, the intervention is no longer seen as a dichotomous variable (subjects either did or did not receive the prevention intervention) but rather as a continuous variable of the degree of implementation. Regression analyses using the degree of implementation not only often provide a more powerful test (under most conditions) of the relationship between intervention and outcome but also can provide some control for alternative hypotheses for results that are based on assuming an all-or-none treatment.

To illustrate, this regression strategy has been used to assess the effects of teachers' instructional practices on students' attitudes and behaviors in a

prevention experiment (Hawkins and Lam 1987). Analyzed at the teacher level, regression analyses revealed important positive links between the use of experimental instructional practices and students' engaged time in class and negative associations with off-task classroom behavior, suggesting that the teaching practices increase classroom involvement as hypothesized. The degree of implementation of the instructional practices also predicted the amount of time students spent on homework, student standardized achievement test scores in math, subject reports of the number of close friends they had at school (an indicator of an hypothesized effect of the cooperative learning methods mentioned earlier), and lower rates of student suspension and expulsion from school as hypothesized. It is important to note that several of these effects did not appear as significant differences between experimental and control groups when analyzed using analysis of covariance that did not include data on the degree of implementation of the instructional practices.

Structural equation modeling methods allow the researcher to further incorporate the intervention implementation data into the analysis of the relationship between the intervention and outcomes. For example, structural equation modeling methods allow (1) estimation of the reliabilities of the measures used to assess the degree of implementation of the intervention, (2) incorporation of reliabilities into the estimation of the independent variables' effects, and (3) the identification and testing of alternative interpretations that could account for the structural linkages between the intervention construct and the outcomes (Costner 1971).

STRUCTURAL EQUATION MODEL REPRESENTATIONS

Figure 1 represents a model tested by the one-way analysis of variance or bivariate regression. This model and its expansions to multiple independent variables or multiple dependent variables assume that the intervention integrity is perfect and without error. Figure 2 introduces several ideas to the model testing process. In this figure, X_2 represents a measure of implementation, and the preventive intervention (X_1) and outcome variable (X_3) are assumed to be operationalizations of the theoretical constructs F_1 and F_2. The model of figure 2 recognizes that the theoretical constructs F_1 and F_2 are not identical with particular operationalizations. The structural model represented in figure 3 extends the model to include multiple indicators (X_2, X_3, X_4) of the theoretical construct (F_1) and represents the implementation and multiple indicators (X_5, X_6, X_7) of the theoretical outcome variable (F_2). Having multiple indicators of the degree of implementation allows the researcher to take into account implementation reliability and test alternative hypotheses for the effects in the study.

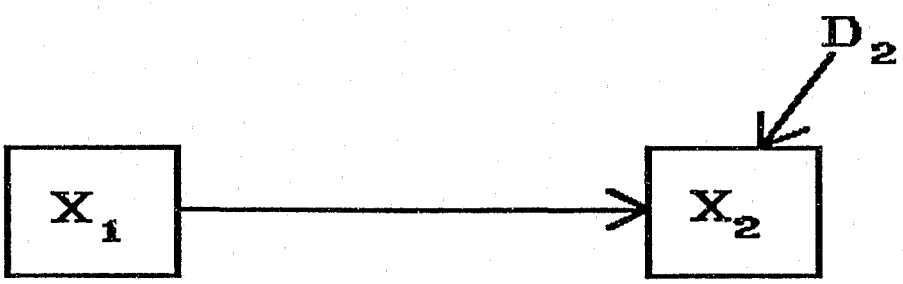

FIGURE 1. *Structural model representation for experiment with a single independent variable manipulated without error (X_1) and a single dependent variable (X_2)*

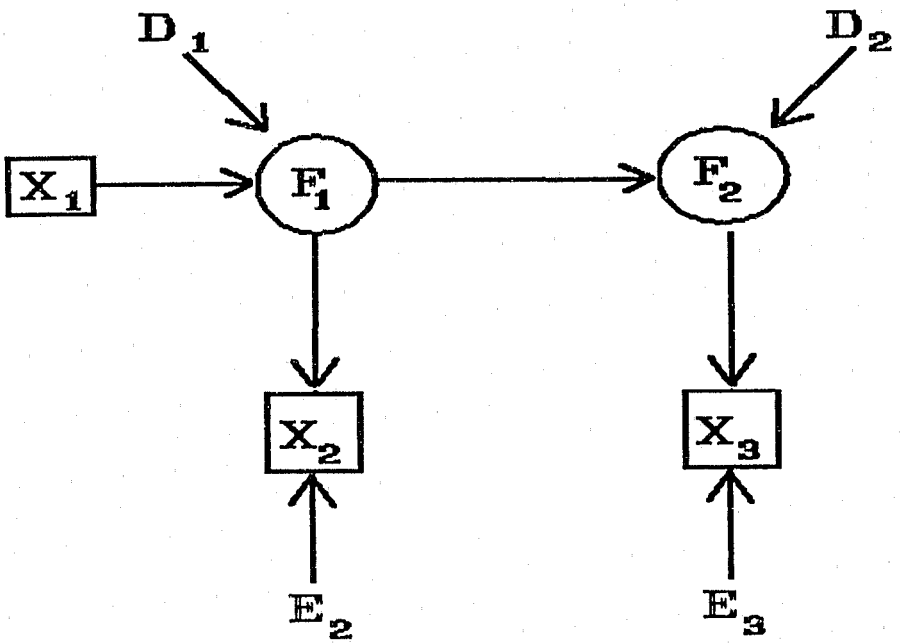

FIGURE 2. *Structural model representation for experiment with single independent variable (F_1), one manipulation check (X_2), one fixed (without error) manipulated manifest variable (X_1), and a single dependent variable (F_2) with one indicator (X_3)*

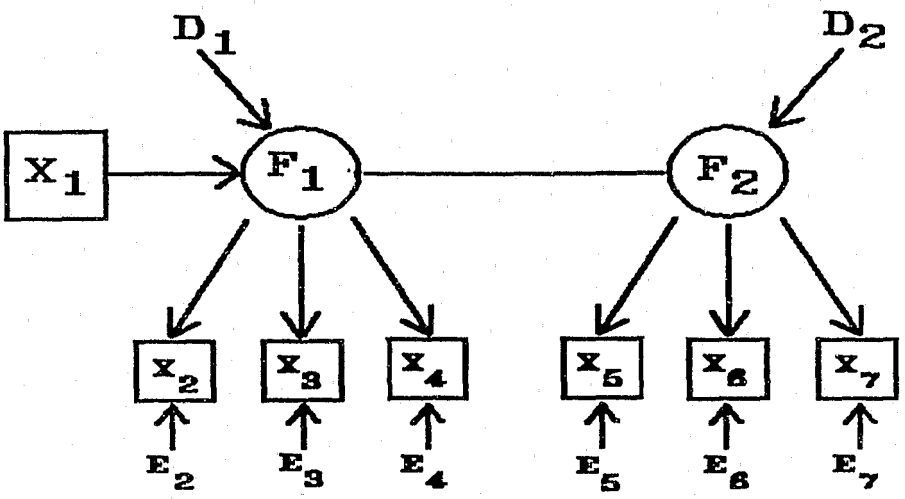

FIGURE 3. *Structural model representation for experiment with single independent variable (F_1), one fixed (without error) manipulated manifest variable (X_1), three manipulation checks (X_2, X_3, X_4), and one dependent (F_2) with three indicators (X_5, X_6, X_7)*

Figure 4 illustrates a model that incorporates the hypotheses that the intervention operationalizes the theoretical construct and shares additional variance with only one of the three implementation measures. This would occur if irrelevant intervention components, such as teacher demand characteristics, also were measured by the teacher implementation (X_2) measures but not by student (X_3) or parent (X_4) measures of intervention implementation. Figure 5 represents a model that extends the pre-post design commonly used in the evaluation of preventive interventions. Using structural equation modeling enables a researcher to account for a correlation between two construct measures (X_3 and X_7) beyond what is explainable on the basis of the theoretical relationship between F_1 and F_2. For example, methods effects due to X_3 and X_7 being measured in the same way can be directly modeled and tested in structural equation approaches.

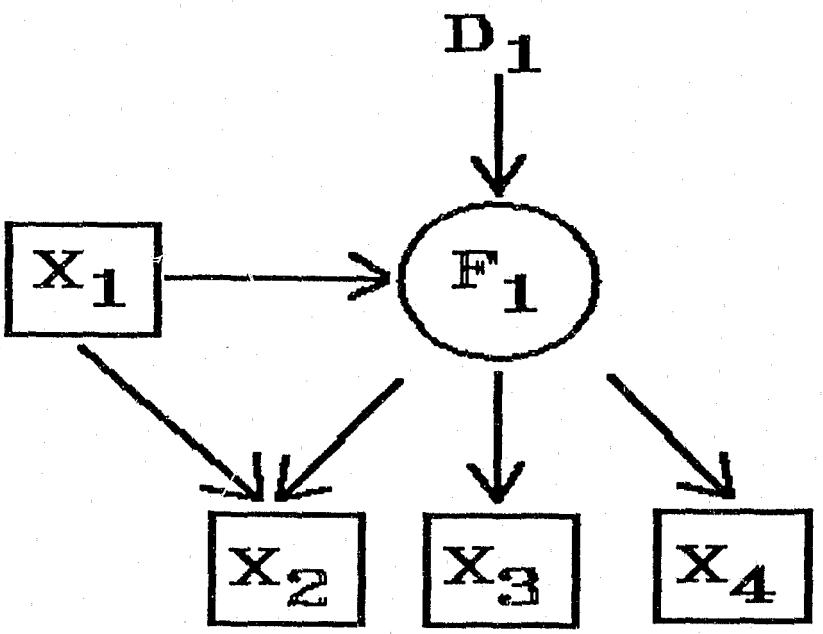

FIGURE 4. *Model for testing an experiment with confounding path (X_1 and X_2)*

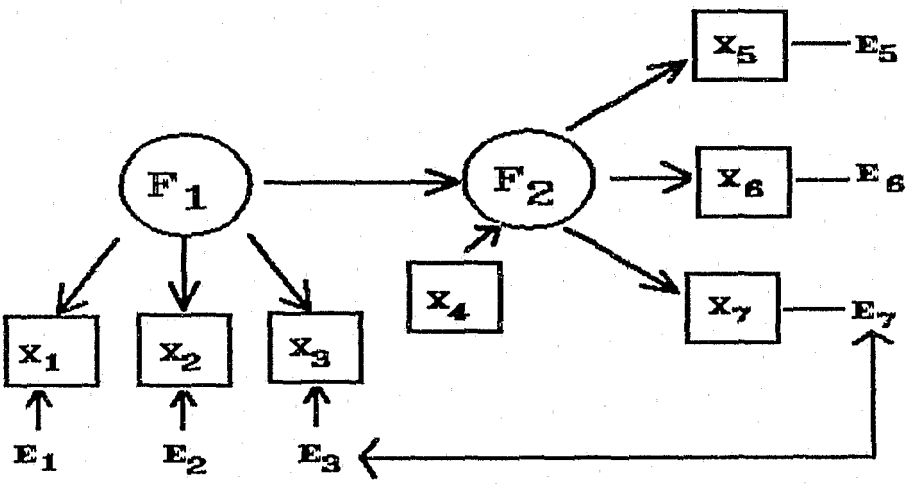

FIGURE 5. *Model representing pretest (F_1), posttest (F_2), and experimental manipulation (X_4)*

Figure 6 represents how a structural equation model can allow a researcher to take into account the reliability of multiple interventions, the short- and long-term outcomes, and the effects of a factor not influenced by the intervention.

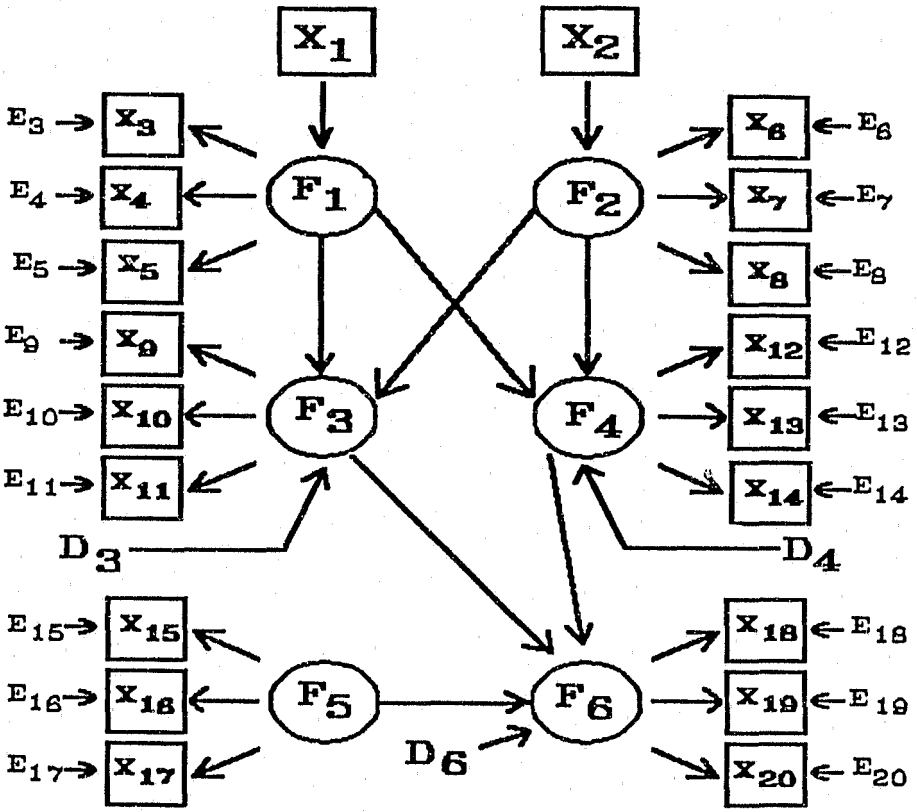

FIGURE 6. *Model representing a study with two intervention operationalized components (X_1 and X_2), two theoretical intervention constructs (F_1 and F_2), each with three implementations measures, two short-term outcomes (F_3 and F_4), one long-term outcome (F_6), and one "other" factor not influenced by the interventions (F_5), which links to the long-term outcome*

SUMMARY

The theory-driven data collection and analysis approach described here implies the need to link proximal intervention outputs (traditionally measured by proportions of subjects initiating use of, occasionally using, or frequently using tobacco, alcohol, or marijuana in groups exposed to different interventions) to more distal outcomes desired such as the prevention of drug abuse that meets psychiatric diagnostic criteria.

This approach requires prospective longitudinal followup studies in which complete panels of subjects who vary with respect to the levels of key predictor constructs are followed up through the period of their highest risk for drug use. Followup studies may need to continue into early adulthood because drug use appears to decline at about age 25 (Elliott et al. 1989).

The conduct of longitudinal panel studies is costly. To justify the cost of longitudinal panel tracking constituted for the evaluation of preventive interventions, more must be learned than whether exposure to a particular intervention was predictive of lower mean levels of drug abuse in the intervention group. Regardless of the answer, the modeling approach discussed here will provide important data to allow for refinement of our understanding of the etiology of drug abuse. Using this approach, nonsignificant differences between groups do not represent failure in long-term evaluation studies, nor do significant differences represent success. Rather, interventions become exogenous variables whose effects on indicators of theoretically relevant predictor constructs and more distal outcomes have been assessed. Hypothesis testing is strengthened to the extent that manipulations of exogenous model variables (interventions) produce hypothesized changes in subsequent endogenous model variables. Where hypothesized relationships are not substantiated, alternative relationships can be modeled and compared with further understanding of the etiology of drug abuse. This activity will allow new understanding to emerge even from those studies in which the interventions fail to prevent drug abuse.

In summary, prevention intervention research should follow the path of nesting preventive interventions within longitudinal panel studies. Well-constructed panel studies of subjects with different but overlapping ages can produce important data on the etiology of drug initiation and abuse and on the effects of developmentally appropriate preventive interventions.

REFERENCES

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-III.* 3d ed. Washington, DC: American Psychiatric Association, 1980.

Biglan, A., and Ary, D.V. Methodological issues in research on smoking prevention. In: Bell, C.S., and Battjes, R., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 170-195.

Botvin, G.J. Substance abuse prevention research: Recent development. *J School Health* 56(9):369-374, 1986.

Botvin, G.J. *Factors Inhibiting Drug Use: Teacher and Peer Effects.* Final Report No. DA-03721, submitted to the National Institute on Drug Abuse, Rockville, MD, 1987.

Catalano, R.F., and Hawkins, J.D. "The Social Development Model: A Theory of Antisocial Behavior." Paper presented at the Safeco Lectureship on Crime and Delinquency, University of Washington at Seattle, 1986.

Costner, H.L. Utilizing causal models to discover flaws in experiments. *Sociometry* 34:398-410, 1971.

Elliott, D.S.; Huizinga, D.; and Menard, S. *Multiple Problem Youth: Delinquency, Substance Use, and Mental Health Problems.* New York: Springer-Verlag, 1989.

Fraser, N.W.; Hawkins, J.D.; and Howard, M.O. Parent training for delinquency prevention: A review. *Child and Youth Services* 11(2):93-125, 1988.

Hall, G.E.; and Loucks, S.F. A developmental model for determining whether the treatment is actually implemented. *Am Educ Res J* 14(3):263-276, 1977.

Hansen, W.B.; Graham, J.W.; Wolkenstein, B.H.; Lundy, B.Z.; Pearson, J.; Flay, B.R.; and Johnson, C.A. Differential impact of three alcohol prevention curricula on hypothesized mediating variables. *J Drug Educ* 18(2):143-153, 1988.

Hawkins, J.D., and Catalano, R.F. *Risk and Protective Factors for Alcohol and Other Drug Problems: Implications for Substance Abuse Prevention.* Seattle: University of Washington, Social Development Research Group, 1989.

Hawkins, J.D.; Catalano, R.F.; and Morrison, D. "The Seattle Social Development Project: Cumulative Effects of Interventions in Grades 1-4." Paper presented at the meeting of the Society for Research in Child Development, Kansas City, MO, April 1989.

Hawkins, J.D.; Doueck, H.; and Lishner, D.M. Changing teaching practices in mainstream classrooms to improve bonding and behavior of low achievers. *Am Educ Res J* 25(1):31-50, 1988.

Hawkins, J.D., and Lam, T. Teacher practices, social development, and delinquency. In: Burchard, J.D., and Burchard, S.N., eds. *The Prevention of Delinquent Behavior.* Beverly Hills: Sage, 1987. pp. 241-274.

Hawkins, J.D.; Lishner, D.M.; and Catalano, R.F. Childhood predictors and the prevention of adolescent substance abuse. In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 75-126.

Johnston, L.D. The etiology and prevention of substance use: What can we learn from recent historical changes? In: Jones, C.L., and Battjes, R.J., eds. *Etiology of Drug Abuse: Implications for Prevention.* National Institute on Drug Abuse Research Monograph 56. DHHS Pub. No. (ADM)87-1335. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 155-177.

Kerr, D.M., and Cummings, C. *Interactive Teaching Map: Classroom Observation Coding Manual.* Seattle: Center for Law and Justice, University of Washington, 1982.

Kerr, D.M.; Kent, L.; and Lam, T.C.M. Measuring program implementation with a classroom observation instrument: The interactive teaching map. *Eval Rev* 9(4):461-482, 1985.

Leinhardt, G. Modeling and measuring educational treatment in evaluation. *Rev Educ Res* 50(3):393-420, 1980.

Leithwood, K.A., and Montgomery, D.J. Evaluating program implementation. *Eval Rev* 4(2):193-214, 1980.

Pentz, M.A.; Dwyer, J.H.; MacKinnon, D.P.; Flay, B.R.; Hansen, W.B.; Wang, E.Y.I.; and Johnson, C.A. A multi-community trial for primary prevention of adolescent drug abuse: Effects on drug use prevalence. *JAMA* 261(22):3259-3266, 1989.

Shaver, J.P. The verification of independent variables in teaching methods research. *Educ Res* 12(10):3-9, 1983.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by National Institute on Drug Abuse grant DA-03721. The authors wish to thank John Campos who performed a literature review and compiled the bibliography for this chapter.

AUTHORS

J. David Hawkins, Ph.D.
Professor and Director

Richard F. Catalano, Ph.D.
Associate Professor

Mary R. Gillmore, Ph.D.
Research Assistant Professor

Social Development Research Group
University of Washington School of Social Work
4101 15th Avenue, N.E., JH-30
Seattle, WA 98195

Robert Abbott, Ph.D.
Professor
Educational Psychology Department
University of Washington
DQ-12
Seattle, WA 98195

140147

Subject Attrition in Prevention Research

Anthony Biglan, Donald Hood, Paul Brozovsky, Linda Ochs, Dennis Ary, and Carol Black

INTRODUCTION

This chapter discusses the role of subject attrition in substance abuse prevention research. Subject attrition routinely occurs in studies designed to evaluate smoking and alcohol and drug abuse prevention programs. Such attrition may affect the validity of experimental comparisons and may limit the extent to which findings can be generalized to adolescents at highest risk. The authors examine concerns about subject attrition, present methods for analyzing attrition in evaluations of prevention programs, and make recommendations for minimizing the extent and impact of attrition in such evaluations.

They also address attrition problems in studies of the prevention of all forms of substance use. However, school-based smoking prevention studies provide the majority of well-controlled research and analysis of attrition issues.

THE EXTENT OF ATTRITION

To estimate the occurrence of attrition in prevention research, we examined smoking prevention and alcohol and drug abuse prevention studies for reports and analyses of attrition. The smoking prevention studies included published studies reviewed by Flay (1985) plus eight studies published since then. The alcohol and drug abuse studies came from a list of evaluations of substance abuse education that Bangert-Drowns (1988) identified as methodologically adequate and providing sufficient information. The studies used to compose these tables are listed in the appendix. Bangert-Drowns (personal communication, April 25, 1989) indicated that he did not exclude studies from his analysis if they failed to report attrition rates but may have eliminated those with high attrition rates or differential attrition between conditions. Thus, the studies he examined probably underestimate the rate of attrition.

Table 1 presents information about the reporting and analysis of subject attrition for these studies. Of the 44 evaluations of smoking prevention programs, 29 (66 percent) reported attrition rates. The mean reported attrition at followup in these studies was 25.8 percent (range 5 percent to 66 percent). Of the 34 alcohol and drug abuse prevention studies, only five (14.7 percent) reported attrition rates. The mean reported rate of attrition at followup was 25.4 percent (range 14 percent to 46 percent). Thus, although many investigators do not report attrition, the available evidence suggests that substantial subject attrition occurs in both types of studies.

TABLE 1. *Reported evaluation of attrition in alcohol and drug abuse prevention studies*

	Smoking Prevention Evaluations (N=44)	Alcohol and Drug Abuse Prevention Studies (N=34)
Studies that reported attrition rates	66.0%	14.7%
Mean rate	25.8%	25.4%
Reported attrition by condition	31.8%	14.7%
Analyzed differences among conditions in:		
Attrition rate	18.2%	0.00%
Substance use	20.4%	0.00%
Analyzed remainder/ dropout difference	22.3%	0.00%

Sources of Attrition

Logic and experience indicate that subject attrition in school-based prevention evaluations is generally due to one of six factors: (1) the student is absent on days when in-school assessments are performed; (2) the student has transferred to another school by the time the postintervention assessment is conducted; (3) the student has dropped out of school by the time of the postintervention assessment; (4) the student fails to come to the assessment (under circumstances in which subjects are taken out of their regular classrooms for assessment); (5) the student declines to continue or the parent declines to have the student continue to participate in the study; or (6) the student "passively" declines, that is, does not complete the questionnaire appropriately.

It also should be noted that subjects drop *in* to studies. For example, in our most recent smoking prevention study, the in-class assessment at 1-year followup found 1,871 subjects who were not previously assessed. This number represented 23 percent of all the subjects we assessed in that year. Depending on the design of the study, the inclusion of these subjects could be important. For example, if substance use prevalence in a school is the dependent variable in the analysis, these subjects should be included. However, if individual subjects are the unit of analysis in a panel design, these subjects cannot be used.

ATTRITION THREATENS INTERNAL VALIDITY

The internal validity of an experimental evaluation is measured by one's confidence that any differences between experimental conditions are due to the experimental variable that was under study rather than to the extraneous factors (Cook and Campbell 1979). For example, in experimentally evaluating a drug abuse prevention program, internal validity is preserved if any observed differences in substance use between those who received the program and those who did not can be attributed to subjects' exposure to the program rather than to other variables such as preexisting differences between conditions in drug use or risk factors for drug use.

Subject attrition can threaten the internal validity of such an experimental comparison. If more subjects drop out of one condition than another, any differences between conditions at postintervention assessments may be simply due to differences in the substance use behavior of subjects who have remained in each experimental condition.

Perhaps the simplest method of controlling for attrition rate differences in experimental conditions is randomization of the units of study (in most cases, schools) to experimental conditions. This method will maximize the equivalence of attrition across groups.

Recommendations for Analysis of Attrition Effects on Internal Validity

Hansen et al. (1985) recommend examining two issues regarding the effects of attrition on internal validity: (1) whether the attrition rate differs among experimental conditions (Hansen et al. 1985) and (2) whether the characteristics of those who remain in the study differ as a function of any of the experimental conditions (Cook and Campbell 1979). However, we believe that the latter issue is more important. Experimental conditions could differ in the rate of attrition, with no difference in substance use rates among those who dropped out of the conditions. In this case, the internal validity of the study

apparently would not be threatened because one would be testing the effects of the intervention on samples that are equivalent on measured variables. Conversely, the conditions could be equivalent in attrition rates, yet have more substance users dropping out of one condition than another. This finding would mean that any outcome differences among conditions could be due to differential loss of substance users, *even though attrition rates were equivalent.* Thus, the only apparent information suggested by significant differences in attrition rate is that group differences on unmeasured variables are possible.

Tests for differences in attrition rate can be provided simply by conducting a chi-square test for the difference between experimental conditions in the proportion of subjects who are missing. However, we advocate a more complete analysis of attrition rate—one that is isomorphic with analysis of the outcome data. For example, if one planned to analyze the effects of a prevention program and its interactions with school grade, gender, and time-1 level of substance use using analysis of variance, the most appropriate analysis of attrition would be an analysis of variance on the proportion of study dropouts with treatment condition, grade, gender, and substance use at initial assessment as the independent variables. This analysis may have more statistical power than one in which only treatment versus control conditions are examined.

Assessment of attrition effects on the characteristics of remaining subjects can be examined by conducting an analysis in which the independent variables are (1) attrition status of subjects at the postintervention time point, (2) treatment condition, and (3) all other independent variables to be included in the analysis of outcome (e.g., gender, age). The inclusion of the other independent variables may contribute to the statistical power of the analysis of this interaction and will allow one to test whether differences in attrition may have influenced any findings of differences on these independent variables. The dependent variables would be measures of subjects' substance use at time 1. A significant interaction between treatment condition and attrition status would indicate that the study dropouts in one condition were significantly higher (or lower) in substance use at time 1 than were the study dropouts in the other condition.

Such differential attrition may compromise the internal validity of the study. For example, if time-1 smokers in the treatment group were more likely to drop out than time-1 smokers in the control group, a finding showing lower smoking at followup among treatment subjects who remained in the study may be due to differential attrition rather than to the intervention.

An alternative approach to the attrition problem that may be useful in some applications has been presented by Allison (1987). Structural equation

modeling techniques using the maximum likelihood method yield estimates of model parameters by comparing subsamples that differ in dropout status. The same methods used to estimate latent variable models are used to estimate models with missing data. Although the method assumes that the data are missing at random (Rubin 1976), the approach may still provide a useful approximation of critical relationships.

Extent of Analysis of Attrition Effects in Prevention Studies

Table 1 presents the percent of smoking or alcohol and drug abuse prevention studies that reported any information about differences in attrition rates across experimental conditions. Of the 44 smoking prevention evaluations, 14 (31.8 percent) reported attrition rates by conditions, and 8 (18.2 percent) reported statistical tests of differences among conditions in these rates. Nine studies (20.4 percent) reported differences among conditions in the substance use patterns of subjects who dropped out versus those who remained in the study, but only four conducted statistical tests to determine whether conditions differed in the attrition rates of substance users. For alcohol and drug abuse studies, 5 (14.7 percent) of the 34 studies reported attrition rates by condition, but none reported statistical tests of differences in rate or interactions between condition and attrition status on measures of substance use.

Evidence That Attrition Has Affected Internal Validity

Of the eight published studies that tested for differential rates of attrition, one found an effect. Of the four studies that examined the interaction of attrition status by condition on substance use measures, one found a significant effect. Biglan et al. (1987a) found no difference in the proportion of subjects who were missing in treatment and control conditions. However, when self-reported smoking rate was used, there was a significant interaction between experimental condition and attrition status at 6 months posttest and an interaction that approached significance ($p<.10$) at 1 year. At both times, more high-rate smokers were missing from the treatment condition than from the control condition. If unanalyzed, this attrition effect would have led to the spurious conclusion that the prevention program had effectively prevented smoking.

Our most recent experimental evaluation of a substance abuse prevention program (described in Biglan et al. 1988 and not included in table 1) has included assessments of treatment and control subjects after 1 and 2 years of intervention. After 1 year, there were no differences between conditions in the proportion of subjects who were missing, nor was there an interaction between attrition status and condition for self-reported smoking or other measures of substance use.

After 2 years, there was no significant difference in attrition rate between treatment and control conditions (chi-square=1.76, p=.19). However, an analysis of variance of attrition rate as recommended above indicated some possibly important differences in attrition associated with the conditions of the experimental design. The analysis treated attrition rate as the dependent variable, with independent variables as follows: treatment versus control, grade (6, 7, 8, or 9 at outset of the study), gender, and smoking status. There was a marginally significant main effect for treatment condition, F(1; 7,738)=2.914, p=.088, and a significant interaction between treatment condition and grade level, F(3; 7,738)=2.97, p=.031. We then analyzed retention separately for each grade and found that for sixth, seventh, and eighth graders there was no difference between treatment and control conditions in the proportion of study dropouts. However, in the ninth grade, treatment subjects (mean=.654) were much more likely to be retained than were control subjects (mean=.562), F(1; 1,897)=12.221, p=.001. Further examination of the data indicated that the effect was caused entirely by one pair of high schools. The control school had a high rate of missing subjects because of the rate of student transfers or dropouts and because of problems we had in obtaining data from absentees. If we drop this pair of schools, there is no difference between treatment and control; if these two schools had been randomly assigned in reverse order, there would have been a marginally significant higher retention rate in the control schools.

The second step in this analysis was to determine if the subjects who remained in the study's treatment condition were different from those who remained in the control condition. The subjects were compared on baseline measures of our primary outcome variables. This analysis showed no condition by retention status interaction for any of the primary variables. Because of the previously reported treatment condition by grade interaction, these variables were analyzed separately for each grade. Because only 1 of the 20 analyses was significant at the .05 level, we concluded that there was no treatment by retention interaction for any of these variables at any grade level. Thus, the dropout rate among ninth graders was higher in the control condition, but those dropping out across treatment and control conditions did not differ on relevant time-1 measures. In this case, the differential attrition appears to be an artifact of differing school attrition rates and seems unrelated to the existence of the intervention in one of the schools.

This study and the one by Biglan et al. (1987a) illustrate the importance of analyzing the interaction of attrition status and condition rather than simply relying on a test of differences among conditions in attrition rate. In the study by Biglan and colleagues (1987a), there was no significant difference in attrition rate, but the loss of smokers was different between treatment and control

conditions. In our more recent study, attrition rates differed, but the time-1 substance use of rates of remaining subjects were equivalent between conditions. Only in the study by Biglan and coworkers (1987a) was internal validity threatened.

ATTRITION THREATENS EXTERNAL VALIDITY

External validity is defined by the degree to which the results of an experimental comparison can be generalized to conditions other than those in which the study was conducted. For example, if a prevention program is found to produce a significant deterrent effect on substance use in one grade level in a particular set of communities, generalizability is the extent to which the same results would occur for different grade levels or in other communities.

Attrition threatens the external validity of prevention research to the extent that subjects who are missing from postintervention evaluations are systematically different from those who remain. In particular, if subjects missing from postintervention assessments were using substances at a higher rate at the preintervention assessment than were subjects who remained, we cannot be sure that any intervention effects demonstrated with the remaining subjects can be generalized to those who are missing.

The effects of attrition on external validity may be somewhat obviated when schools are used as the unit of analysis in outcome studies. It has been noted elsewhere (Biglan and Ary 1985) that the ideal method of analyzing the effects of prevention programs is to use schools rather than individual subjects as the unit of analysis (e.g., Biglan and Ary 1985). Where possible, this method provides the opportunity to include data from all of the subjects rather than only those for whom data were obtained at both pretreatment and postintervention assessments. Thus, subjects who have dropped *in* to the study can be included. Because, as presented below, these subjects are more like dropouts than those who remain, such an approach should increase the external validity of the study.

Recommendations for Analysis of Effects of Attrition on External Validity

Hansen and colleagues (1985) recommend that the effects of attrition on external validity be evaluated in two ways. First, the pretest scores for dropouts and remaining subjects can be compared to see if those who remain are different from those who are missing. Second, if data are collected on more than two occasions, one can examine whether the subjects who are missing at time 3 (or beyond) differ from remaining subjects on change scores from time 1 to time 2. Inclusion of missing data using the structural equation modeling

methods outlined by Allison (1987) represents another approach to the problem of attrition and generalizability.

Extent of Analysis of Differences Between Study Dropouts and Remaining Subjects

As table 1 indicates, only 22.3 percent of the smoking prevention studies have examined differences between study dropouts and remaining subjects in substance use or other characteristics, and *none* of the alcohol and drug abuse prevention studies has done so.

Evidence That Those Who Are Missing Are Systematically Different From Those Who Remain

The evidence that study dropouts are systematically different from those who remain is quite strong. The strongest evidence comes from a study by Pirie and coworkers (1988) in which study dropouts were tracked and assessed. Pirie and colleagues returned to the schools 10 days after initial assessment to get data from absentees and used telephone tracking procedures to find subjects who were no longer in the school district. They were able to obtain data from 87.6 percent of the subjects who had entered the study 5 or 6 years previously. This included 90.5 percent of the subjects who were no longer in the district. They classified subjects into four categories: (1) in school on the day of assessment, (2) absent on the day of assessment, (3) transferred to a different school, and (4) dropped out of school. The prevalence of daily smoking among those who had dropped out was substantially and significantly higher than for the other groups (77.7 percent for the dropouts versus 19.3 percent for those who were in school on the day of assessment). Smoking prevalence among transfer students also was significantly higher than that for subjects who were in school on the day of assessment. These data indicate quite clearly that subjects who are typically missing from smoking prevention studies are more likely to be smokers than those who remain. Given the well-established relationship between smoking and the use of other substances (Osgood et al. 1988), the results also suggest that smoking prevention study dropouts are users of other substances.

Although the majority of studies have reported no analyses of this issue, each of those that did report such analyses found that the subjects who were missing from postintervention assessment were systematically different from those who remained. Those who were missing had reported significantly more smoking and use of other substances at time 1 (Biglan et al. 1987a, 1987b; Flay et al. 1987; Flay et al., in press; Johnson et al. 1986; Hansen et al. 1985, 1988; Murray et al. 1987). Evidence also shows that study dropouts have more

people in their environment who smoke (parents, friends, and brothers), indicate greater intentions to smoke, are lower in educational aspirations, have less well-educated parents, and receive more offers of cigarettes than those who remain in the study (Biglan et al. 1987a, 1987b).

Ellickson and colleagues (1988) reported that eighth-grade subjects who had transferred to another school differed from those who remained in the same school. Those who had transferred were more likely to be from a minority group, to have a disrupted family, to have parents who had not completed high school, to have grades of C or lower, to have been absent frequently, to report deviant behavior, and to have used cigarettes, marijuana, or alcohol.

The Variables Discriminating Study Dropouts From Remaining Subjects

Discriminant analysis provides a more complete and informative analysis of the differences between study dropouts and those who remained. We conducted such analyses on the data from our most recent smoking prevention study. We first looked at differences between subjects who remained at 1-year followup and subjects who were missing by that time. We then tested the replicability of this analysis by looking at differences between subjects who were present at 1-year and 2-year followups and those who were present at 1-year followup but missing at 2-year followup. Table 2 presents the variables that discriminated those who remained and study dropouts at 1 year and the variables that did so at 2-year followup.

Despite the rather large differences between these groups on many measures, the percent of variance accounted for and the accuracy in prediction of group membership were not high. At year 1, 13 variables contributed significantly to discrimination between these groups, $F(13; 7,404)=35.17$, $p<.001$. The function accounted for 6 percent of the variance. The function correctly identified 97.7 percent of those who remained, but only 10.3 percent of the study dropouts. At year 2, eleven variables discriminated the groups, $F(11; 8,061)=33.74$, $p<.001$. Nine of these variables were the same ones that entered the discriminant function in the year-1 analysis. The year-2 function accounted for 4.6 percent of variance; it correctly identified 97.6 percent of those who remained but only 8.6 percent of the study dropouts.

Implications

Although the evidence reviewed here clearly shows that prevention study dropouts are more likely than those who remained to be substance users, the precise implications of this fact for the external validity of prevention evaluations should be noted. Strictly speaking, this evidence does not necessarily imply

that prevention programs found to deter substance use among remaining subjects did not deter substance use among those who were lost to followup. Had these subjects been found, we may have detected a deterrent program effect among them as well. The best evidence we have in this regard comes from the study by Pirie and coworkers (1988). The high rate of smoking that they found among subjects who would have been missing in most prevention

TABLE 2. *Variables significantly discriminating study dropouts from remaining subjects*

	Year-1 Analysis		Year-2 Analysis	
Variable	Univariate F (df=1; 7,416)	F to Remove*	Univariate F (df=1; 8,071)	F to Remove*
Addicted smoking	251.40	18.06	192.90	7.46
Father's education	66.56	21.44	47.95	8.14
Age	79.17	18.34	63.62	23.33
Mother's smoking	81.93	19.83	74.69	15.97
Ethnic self-description	25.61	18.42	17.85	12.47
Expired air CO	157.40	19.93	138.20	25.27
Friends' smoking	221.60	24.47	179.60	17.61
Daily marijuana smoking	90.23	7.35	45.85	
Daily alcohol consumption	7.11	6.32	24.89	
Seatbelt use	65.28	4.09	86.77	14.03
Father's smokeless use	0.00	3.87	.17	3.79
Grade level	47.33		35.90	
Probability smoking in 1 year	167.70		105.60	
Alcohol use index	17.29		25.00	
Marijuana use index	56.07		28.31	
Mother's education	55.09		37.29	
Hard drug use	120.90		53.64	
Brother's smoking	45.92		43.89	
Best friend's smoking	151.30		179.60	
Gender	6.75	13.86	.01	
Smoking index	75.02		90.82	
Chewing index	9.16		29.15	11.69
Father's smoking	57.36		31.10	
Probability chewing in 1 year	19.21		9.92	6.05
Sister's smoking	36.38		37.63	
Friends' smokeless use	28.09	6.77	29.29	

*Only significant Fs are shown.

studies suggests that most of these subjects were smokers. However, they did not analyze for differences among conditions for this subgroup (B. Murray, personal communication, April 1989). Although we think it unlikely, such an analysis would have shown that treatment deterred smoking among these "high-risk" subjects. Biglan and colleagues (1987a, 1987b) and Ary and coworkers (1989) have presented evidence of intervention effects among smokers who remained in the study.

The evidence on subject attrition indicates that estimates of preventive effects derived from prevention studies probably overestimate the degree to which the population prevalence of substance use is being reduced. Even if intervention effects are generalizable to dropouts, the overall prevalence of smoking in treatment schools is higher than assessments of remaining subjects would suggest. It may well be that the prevalence of smoking among dropouts is not reduced.

THE VALUE OF INCLUDING STUDY "DROP-INS"

As noted above, subjects also drop *in* to prevention studies. These subjects tend to be more like those who drop out of studies than like those who remain. For example, for our most recent smoking prevention study, we conducted a discriminant analysis of year-1 drop-ins vs. subjects who had been in the study at pretreatment intervention. Ten variables contributed significantly to discrimination between these groups, $F(10; 8,062)=28.41$, $p<.001$. They were expired air CO, alcohol use, mother's education, age, friends' smoking, ethnic identity, mother's smoking, friends' use of smokeless tobacco, father's use of smokeless tobacco, and father's education. Seven of these variables also discriminated those who remained from study dropouts in the analyses described above.

These results underscore the value of including study drop-ins when subjects are not used as the unit of analysis. Their inclusion will increase the generalizability of the study and could reduce differences in attrition between treatment conditions. (One may want to obtain data on when these subjects entered the school to ensure that only those who actually received the intervention are included in the treatment schools and that an equivalent group of subjects is included in the control schools.)

EFFICACY AND EFFECTIVENESS TRIALS

Flay (1986) has suggested the value of distinguishing between efficacy and effectiveness trials. In the former studies, the effects of a prevention program are evaluated under optimal conditions, including random assignment of

subjects to conditions, uniform delivery of the intervention to a specified target audience, and optimal acceptance and participation in the program by the recipients. Effectiveness trials are concerned with determining whether a program does more harm than good when it is delivered in circumstances that are likely to lead some members of the target audience not to receive the program and/or some persons not to accept or participate in the program.

Some evidence from analysis of attrition in prevention studies is relevant to this distinction. Analyses of the attrition problem have led us to compare the characteristics of absentees with those of subjects who are available for assessment the first time we go to each school. As noted above (and in Biglan et al. 1987b), we have found that absentees have higher rates of substance use than present subjects. To the extent that absence is not due to our conducting the assessment, this suggests that evaluations of school-based prevention programs generally fail even to reach some of the higher risk subjects. The evaluations are thus more effectiveness trials than efficacy trials. As Flay (1985) points out, any failures of the evaluated programs to affect substance use may be due to program inefficacy, but they also could be caused by failures of efficacious programs to reach some target subjects.

METHODS OF REDUCING ATTRITION

Procedures To Attain the Largest Possible Proportion of Students

Two things have helped us to find students in subsequent years. First, we routinely check the names of students on mailing labels (which we get for the purpose of mailing statements of informed consent to homes) against class rosters. This procedure allows us to identify students who are not in the classes in which we intend to conduct our assessments. Second, we verify information from teachers concerning student withdrawal from class or school with the school records department. This verification sometimes lets us find students who otherwise would have been assumed to have left the school.

We also have found that notifying teachers and administrators 1 week in advance of assessments as to which classes we will be going into and which students we will be assessing has helped us to obtain teacher cooperation. This method avoids scheduling assessments at times that conflict with tests or other classwork the students cannot miss. It also has provided us with another source of information as to where we might locate a given student (e.g., "He never comes on Monday but is usually here on Tuesday."). Another way we have increased teacher and administrative cooperation is to schedule assessments in the first week of the semester or the week before quarter or semester exams or during exam week. In addition, to delineate the cause of

attrition, we negotiate access to school records that indicate what happened to the student.

Returning for Absentees

Those who are absent on the day of assessment are more likely to be substance users than those who are present (Ellickson et al. 1988). Thus, the simplest and least expensive method of reducing attrition among substance-using students is to minimize the number of absentees. We try to avoid scheduling assessments on Mondays and Fridays or at any time close to vacations. Within 2 weeks of the initial assessment, we return to the school to obtain data from students who were absent on the day of assessment. In high schools or large middle schools, we go directly to the individual classroom, as opposed to having the teacher send the student(s) to the designated place. We escort the students to a central location to complete the assessment. In very small high schools and small middle schools, we have been successful in having the students who were absent during the initial assessment called to a central location at the beginning of a designated class period. This arrangement is made through school administrators, and the information is sent to teachers, usually via a daily bulletin. If students are not in the classes that we had expected them to be, we go back to the school records to verify that they are still in school and to find them in other classes. Sometimes teachers are reluctant to let students out of classes; in these cases, we negotiate with the students to complete the assessment in a different period.

In 3 years of followup assessments of subjects in our most recent prevention study (Ary et al. 1989), the following proportion of students were on a class roster but absent on the day of assessment: 8.2 percent for 1-year followup, 9.9 percent for 2-year followup, and 12.7 percent for 3-year followup. As we have developed the procedures just described for getting data from these students, our success has improved: The percent of absentees from whom data were obtained was 54 percent and 50 percent, respectively, in the first two followups, and 70 percent in the most recent assessment.

Tracking Those Who Are No Longer in School

Some students who are missing from the school in which they were originally assessed can be located in other schools that are participating in the study. Ellickson and coworkers (1988) report a method of tracking subjects who are no longer in a study school. Among a group of 1,045 eighth-grade students in this category, they were able to locate 77 percent and obtain data from 66 percent of them. Their procedures involved (1) sending a questionnaire to the school to which the student had transferred when this information was available or (2)

mailing to their old home address a request for change of address. Students were offered $5 for completing the questionnaire. Thirty-four percent of the students from whom they obtained data were contacted through the mailing to the schools; this procedure is an important adjunct to efforts to reach students through mailings to their homes. The procedure might be less effective among older cohorts of subjects because a larger percentage of them would presumably have dropped out of school.

Pirie and colleagues (1988) reported methods of tracking dropouts and transfers in a cohort of 7,124 students who had been in seventh grade at the outset of the study and were being followed up 5 or 6 years later. The subjects were located on the basis of information about them and their families that had been obtained in previous years. Further efforts to find these subjects involved the use of telephone directories, mailing requests to old addresses for address corrections, and calling people with the same last name. They were able to locate and interview by telephone 90 percent of the 1,551 subjects who otherwise would have been lost.

Through in-school surveys and telephone tracking procedures, Pirie's group was able to obtain data from 87.6 percent of the original sample. It should be noted, however, that their sample had a fairly low school dropout rate (4.4 percent of the measured sample).

How Important Is It To Track Missing Subjects?

Procedures for tracking missing subjects may reduce the probability of differential attrition among experimental conditions. Such tracking also will increase the generalizability of the study results by including more subjects who are substance users. However, it should be noted that, aside from returning to assess absentees, these procedures can be quite expensive. What is needed is a cost/benefit analysis of the value of such procedures for improving the validity of prevention research. With respect to internal validity, analyses of the likelihood that these procedures reduce differential attrition are needed. With respect to attrition's threat to external validity, analyses are needed of the degree to which treatment effects are the same or different for those who can be assessed only by tracking beyond their original study school. In other words, if a prevention study shows effects on those who are assessed in their original school but not among those who are missing from in-school assessments (assuming adequate statistical power), the case for tracking of this sort will be strengthened. If, on the other hand, a series of studies indicate that treatment effects are the same for tracked and original-school subjects, the need for tracking would appear less valid.

We recommend that the next wave of prevention evaluations be designed to track subjects who are missing from their original schools and to allow determination of the value of such tracking in increasing the internal and external validity of these studies.

SUMMARY

Subject attrition threatens the internal validity of substance abuse prevention studies because differences in the rate of attrition and the substance use behavior of remaining subjects in the different conditions could account for any differences found in substance use rates. Attrition threatens the external validity of prevention studies because, to the extent that study dropouts are different from remaining subjects, the results of the study may not be generalizable to study dropouts. Analysis of these threats to the validity of prevention studies should be routinely conducted. However, studies of alcohol and drug abuse prevention have generally failed to report or analyze subject attrition. Smoking prevention studies have more frequently reported attrition, and they have recently begun to analyze the degree to which attrition may affect the internal and external validity of the study. Evidence thus far suggests that differences in attrition across conditions do occur occasionally. The evidence is substantial that study dropouts are systematically more likely to smoke, to use other substances, and to score highly on other risk-taking measures.

REFERENCES

Allison, P.D. Estimation of linear models with incomplete data. *Sociol Methodol 1987* 17:71-103, 1987.

Ary, D.V.; Biglan, A.; Glasgow, R.; Zoref, L.; Black, C.; Ochs, L.; Severson, H.; Kelly, R.; Weissman, W.; Lichtenstein, E.; Brozovsky, P.; Wirt, R.; and James, L. The efficacy of social-influence prevention programs versus standard care: Are new initiatives needed? Submitted for publication, 1989.

Bangert-Drowns, R.L. The effects of school-based substance abuse education—a meta-analysis. *J Drug Educ* 18(3):243-264, 1988.

Biglan, A., and Ary, D.V. Methodological issues in research on smoking prevention. In: Bell, C.S., and Battjes, R.J., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 170-195.

Biglan, A.; Glasgow, R.; Ary, D.; Thompson, R.; Severson, H.; Lichtenstein, E.; Weissman, W.; Faller, C.; and Gallison, C. How generalizable are the effects of smoking prevention program? Refusal skills training and parent messages in a teacher-administered program. *J Behav Med* 10(6):613-628, 1987a.

Biglan, A.; James, L.E.; LaChance, P.; Zoref, L.; and Joffe, J. Videotaped materials in a school-based smoking prevention program. *Prev Med* 17:559-584, 1988.

Biglan, A.; Severson, H.; Ary, D.; Faller, C.; Gallison, C.; Thompson, R.; Glasgow, R.; and Lichtenstein, E. Do smoking prevention programs really work? Attrition and the internal and external validity of an evaluation of a refusal skills training program. *J Behav Med* 10(2):159,1987b.

Cook, T.D., and Campbell, D.T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally, 1979.

Ellickson, P.L.; Bianca, D.; and Schoeff, D.C. Containing attrition in school-based research. *Eval Rev* 12(4):331-351, 1988.

Flay, B.R. Psychosocial approaches to smoking prevention: A review of findings. *Health Psychol* 4(5):449-488, 1985.

Flay, B.R. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 15:451-474, 1986.

Flay, B.R.; Hansen, W.B.; Johnson, C.A.; Collins, L.M.; Dent, C.W.; Dwyer, K.M.; Grossman, L.; Hockstein, G.; Rauch, J.; Sobel, J.L.; Sobol, D.F.; Sussman, S.; and Ulene, A. Implementation effectiveness trial of a social influences smoking prevention program using school and television. *Health Educ Res* 2(4):385-400, 1987.

Flay, B.R.; Koepke, D.; Thompson, S.; Santi, S.; Best, J.A.; and Brown, S.K. Six-year follow-up of the first Waterloo school smoking prevention trial. *Am J Public Health* 79(10):1371-1376, 1989.

Hansen, W.B.; Collins, L.M.; Malotte, C.K.; Johnson, C.A.; and Fielding, J.E. Attrition in prevention research. *J Behav Med* 8(3):261-275, 1985.

Hansen, W.B.; Johnson, C.A.; Flay, B.R.; Graham, J.W.; and Sobel, J. Affective and social influences approaches to the prevention of multiple substance abuse among seventh grade students: Results from project SMART. *Prev Med* 17(2):135-154, 1988.

Johnson, C.A.; Hansen, W.B.; Collins, L.M.; and Graham, J.W. High school smoking prevention: Results of a three-year longitudinal study. *J Behav Med* 9(5):439, 1986.

Murray, D.M.; Richards, P.S.; Luepker, R.V.; and Johnson, C.A. The prevention of cigarette smoking in children: Two- and three-year followup comparisons of four prevention strategies. *J Behav Med* 10(6):595-611, 1987.

Osgood, D.W.; Johnston, L.D.; O'Malley, P.M.; and Bachman, J.G. The generality of deviance in late adolescence and early adulthood. *Am Sociol Rev* 53:81-93, 1988.

Pirie, P.L.; Murray, D.M.; and Luepker, R.V. Smoking prevalence in a cohort of adolescents, including absentees, dropouts, and transfers. *Am J Public Health* 78:176-178, 1988.

Rubin, D.B. Inference and missing data. *Biometrika* 63:581-592, 1976.

APPENDIX

Smoking Prevention Studies

Arkin, R.M.; Roemhild, H.F.; Johnson, C.A.; Luepker, R.V.; and Murray, D.M. The Minnesota smoking prevention program: A seventh-grade health curriculum supplement. *J Sch Health* 51(9):611-616, 1981.

Best, J.A.; Flay, B.R.; Towson, S.M.J.; Ryan, K.B.; Perry, C.L.; Brown, K.S.; Kersell, M.W.; and D'Avernas, J.R. Smoking prevention and the concept of risk. *J Appl Soc Psychol* 14(3):257-273, 1984.

Biglan, A.; Glasgow, R.; Ary, D.; Thompson, R.; Severson, H.; Lichtenstein, E.; Weissman, W.; Faller, C.; and Gallison, C. How generalizable are the effects of smoking prevention programs? Refusal skills training and parent messages in a teacher-administered program. *J Behav Med* 10(6):613-628, 1987.

Biglan, A.; Severson, H.; Ary, D.; Faller, C.; Gallison, C.; Thompson, R.; Glasgow, R.; and Lichtenstein, E. Do smoking prevention programs really work? Attrition and the internal and external validity of an evaluation of a refusal skills training program. *J Behav Med* 10(2):159, 1987.

Botvin, G.J.; Baker, E.; Renick, N.L.; Filazzola, A.D.; and Botvin, E.M. A cognitive-behavioral approach to substance abuse prevention. *Addict Behav* 9:137-147, 1984.

Botvin, G.J., and Eng, A. A comprehensive school-based smoking prevention program. *J Sch Health* 50(4):209-213, 1980.

Botvin, G.J., and Eng, A. The efficacy of a multicomponent approach to the prevention of cigarette smoking. *Prev Med* 11:199-211, 1982.

Botvin, G.J.; Eng, A.; and Williams, C.L. Preventing the onset of cigarette smoking through life skills training. *Prev Med* 9:135-143, 1980.

Botvin, G.J.; Renick, N.; and Baker, E. The effects of scheduling format and booster sessions on a broad-spectrum psychosocial approach to smoking prevention. *J Behav Med* 6(4):359-379, 1983.

Evans, R.E.; Rozelle, R.M.; Maxwell, S.E.; Raines, B.E.; Dill, C.A.; Guthrie, T.J.; Henderson, A.H.; and Hill, P.C. Social modeling films to deter smoking in adolescents: Results of a three-year field investigation. *J Appl Psychol* 66:399-414, 1981.

Evans, R.E.; Rozelle, R.M.; Mittelmark, M.; Hansen, W.B.; Bane, A.; and Havis, J. Deterring the onset of smoking in children: Knowledge of immediate physiological effects and coping with peer pressure, media pressure, and parent modeling. *J Appl Soc Psychol* 8(2):1260135, 1978.

Flay, B.R.; D'Avernas, J.R.; Best, J.A.; Kersell, M.W.; and Ryan, K.B. Cigarette smoking: Why young people do it and ways of preventing it. In: McGrath, P., and Firestone, P., eds. *Pediatric and Adolescent Behavioral Medicine.* New York: Springer-Verlag, 1983.

Flay, B.R.; Hansen, W.B.; Johnson, C.A.; Collins, L.M.; Dent, C.W.; Dwyer, K.M.; Grossman, L.; Hockstein, G.; Rauch, J.; Sobel, J.L.; Sobol, D.F.; Sussman, S.; and Ulene, A. Implementation effectiveness trial of a social influences smoking prevention program using school and television. *Health Educ Res* 2(4):385-400, 1987.

Flay, B.R.; Koepke, D.; Thompson, S.; Santi, S.; Best, J.A.; and Brown, S.K. Six-year follow-up of the first Waterloo school smoking prevention trial. *Am J Public Health* 79(10)1371-1376, 1989.

Flay, B.R.; Ryan, K.B.; Best, A.; Brown, S.; Kersell, M.W.; D'Avernas, J.R.; and Zanna, M.P. Are social-psychological smoking prevention programs effective? The Waterloo study. *J Behav Med* 8(1):37-59, 1985.

Gilchrist, L.D., and Schinke, S.P. Self-control skills for smoking prevention. *Advances in Cancer Control: Epidemiology and Research.* New York: Alan R. Liss, Inc., 1984.

Gilchrist, L.D.; Schinke, S.P.; and Blythe, B.J. Primary prevention services for children and youth. *Children Youth Serv Rev* 1:379-391, 1979.

Hansen, W.B.; Collins, L.M.; Malotie, C.K.; Johnson, C.A.; and Fielding, J.E. Attrition in prevention research. *J Behav Med* 8(3):261-275, 1985.

Hansen, W.B.; Johnson, C.A.; Flay, B.R.; Graham, J.W.; and Sobel, J. Affective and social influences approaches to the prevention of multiple substance abuse among seventh grade students: Results from Project SMART. *Prev Med* 17(2):135-154, 1988.

Hurd, P.D.; Johnson, C.A.; Pechacek, T.; Blast, L.P.; Jacobs, D.R.; and Luepker, R.V. Prevention of cigarette smoking in seventh grade students. *J Behav Med* 3(1):15-28, 1980.

Luepker, R.V.; Johnson, C.A.; Murray, D.M.; and Pechacek, T.F. Prevention of cigarette smoking: Three-year followup of an education program for youth. *J Behav Med* 6(1):53-62, 1983.

McAlister, A., and Gordon, N.P. Attrition bias in a cohort study of substance abuse onset and prevention. *Eval Rev* 10:853-859, 1987.

McAlister, A.; Perry, C.; Killen, J.; Slinkard, L.A.; and Maccoby, N. Pilot study of smoking, alcohol and drug abuse prevention. *Am J Public Health* 70(7):719-721, 1980.

McAlister, A.L.; Perry, C.; and Maccoby, N. Adolescent smoking: Onset and prevention. *Pediatrics* 63(4):650-658, 1979.

Murray, D.M.; Johnson, C.A.; Luepker, R.V.; and Mittelmark, M.B. The prevention of cigarette smoking in children: A comparison of four strategies. *J Appl Soc Psychol* 14(3):274-288, 1984.

Murray, D.M.; Pirie, P.; Luekpker, R.V.; and Pallonen, U. Five- and six-year followup results from four seventh-grade smoking prevention strategies. *J Behav Med,* in press.

Murray, D.M.; Richards, P.S.; Luepker, R.V.; and Johnson, C.A. The prevention of cigarette smoking in children: Two- and three-year followup comparisons of four prevention strategies. *J Behav Med* 10(6):595-611, 1987.

Pentz, M.A. Prevention of adolescent substance abuse through social skill development. In: Glynn, T.J.; Leukefeld, C.G.; and Ludford, J.P., eds. *Preventing Adolescent Drug Abuse: Intervention Strategies.* National Institute on Drug Abuse Research Monograph 47. DHHS Pub. No. (ADM)83-1280. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1983. pp. 195-232.

Pentz, M.A. Social competence and self-efficacy as determinants of substance use in adolescence. In: Shiffman, S., and Wills, T.A., eds. *Coping and Substance Use.* New York: Academic Press, 1985. pp. 117-142.

Perry, C.; Killen, J.; Telch, M.; Slinkard, L.A.; and Danaher, B.G. Modifying smoking behavior of teenagers: A school-based intervention. *Am J Public Health* 70(7):722-725, 1980b.

Perry, C.L.; Killen, J.; and Slinkard, L.A. Peer teaching and smoking prevention among junior high students. *Adolescence* 15(58):277-281, 1980a.

Perry, C.L.; Telch, M.J.; Killen, J.; Burke, A.; and Maccoby, N. High school smoking prevention: The relative efficacy of varied treatments and instructions. *Adolescence* 28(71):561-566, 1983.

Schinke, S.P., and Blythe, B.J. Cognitive-behavioral prevention of children's smoking. *Child Behav Ther* 3(4):25-42, 1981.

Schinke, S.P., and Gilchrist, L.D. Primary prevention of tobacco smoking. *J Sch Health* 53(7):416-419, 1983.

Schinke, S.P., and Gilchrist, L.D. Preventing cigarette smoking with youth. *J Primary Prev* 5(1):48-56, 1984.

Telch, M.J.; Killen, J.D.; McAlister, A.L.; Perry, C.L.; and Maccoby, N. Long-term followup of a pilot project on smoking prevention with adolescents. *J Behav Med* 5(1):1-8, 1982.

Tell, G.S.; Klepp, K.I.; Vellar, O.D.; and McAlister, A. Preventing the onset of cigarette smoking in Norwegian adolescents: The Oslo youth study. *Prev Med* 13:256-275, 1984.

Vartiainen, E.; Pallonen, U.; McAlister, A.; Koskela, K.; and Puska, P. Four-year followup results of the smoking prevention program in the North Karelia youth project. *Prev Med* 15:692-698, 1986.

Wills, T. Stress, coping, and tobacco and alcohol use in early adolescence. In: Shiffman, S., and Wills, T.A., eds. *Coping and Substance Use.* New York: Academic Press, 1985. pp. 63-94.

Alcohol and Drug Abuse Prevention Studies

Beal, L.E. Youth education about alcohol: Evaluation of a multilevel program conducted by teachers and peer leaders. *Dissertation Abstr Int* 38:7161A-7162A, 1977.

Benberg, T.E. The effects of a planned curriculum on correlates of drug-abusing behavior. *Dissertation Abstr Int* 34:7095A-7096A, 1973.

Dielman, T.E.; Shope, J.T.; Butchart, A.T.; and Campanelli, P.C. Prevention of adolescent alcohol misuse: An elementary school program. *J Pediatr Psychol* 11(2):259-282, 1986.

Duryea, E.J. Utilizing tenets of inoculation theory to develop and evaluate a preventive alcohol education intervention. *J Sch Health* 53(4):250-256, 1983.

Duryea, E.J.; Mohr, P.; Newman, I.M.; Martin, G.L.; and Egwaoje, E. Six-month followup results of a preventive alcohol education intervention. *J Drug Educ* 14(2):97-104, 1984.

English, G.E. The effectiveness of emotional-appeal versus fact-giving drug educational films. *J Sch Health* 42(9):540-541, 1972.

Engs, R.C.; DeCoster, D.; Larson, R.V.; and McPheron, P. The drinking behavior of college students and cognitive effects of a voluntary alcohol education program. *Nat Assoc Personnel Administrators J* 15:59-63, 1978.

Friedman, S.M. A drug education program emphasizing affective approaches and its influence upon intermediate school student and teachers attitudes. *Dissertation Abstr Int* 34:2270A, 1973.

Goodstadt, M.S.; Sheppard, M.A.; and Chang, G.C. An evaluation of two school-based alcohol education programs. *J Stud Alcohol* 43(3):352-369, 1982.

Jackson, J.A. An evaluation of the students, teachers, and residents involved in drug education (STRIDE) program. *Dissertation Abstr Int* 36:6357B, 1975.

Johnson, C.A.; Hansen, W.B.; Collins, L.M.; and Graham, J.W. High-school smoking prevention: Results of a three-year longitudinal study. *J Behav Med* 9(5):439, 1986.

Kearney, A.L., and Hines, M.H. Evaluation of the effectiveness of a drug prevention education program. *J Drug Educ* 10(2):127-134, 1980.

Lewis, J.M.; Gossett, J.T.; and Phillips, V.A. Evaluation of a drug prevention program. *Hosp Community Psychiatry* 23:124-126, 1972.

Mascoll, S.H. The effect of alcohol education programs on the knowledge, attitudes, and intended behavior of eighth grade students. *Dissertation Abstr Int* 37:2003A-2004A, 1976.

McAlister, A.; Perry, C.; Killen, J.; Slinkard, L.A.; and Maccoby, N. Pilot study of smoking, alcohol, and drug abuse prevention. *Public Health Beliefs* 70(7):719-721, 1980.

Myers, E.E. The effects of a drug education curriculum based on a causal approach to human behavior. *J Drug Educ* 4(3):309-315, 1974.

National Institute on Drug Abuse. *The NAPA Drug Abuse Prevention Project: Research Findings.* DHHS Pub. No. (ADM)84-1339. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1984.

Newman, I.M.; Mohr, P.; Badger, G.; and Gillespie, T.S. Effects of teacher preparation and student age on an alcohol and drug education curriculum. *J Drug Educ* 14(1):23-35, 1984.

O'Rourke, T.W. Assessment of the effectiveness of the New York State drug curriculum guide with respect to drug knowledge. *J Drug Educ* 3(1):57-66, 1973.

O'Rourke, T.W., and Barr, S.L. Assessment of the effectiveness of the New York State drug curriculum guide with respect to drug attitudes. *J Drug Educ* 4(3):347-356, 1974.

Perry, C.; Killen, J.; Telch, M.; Slinkard, L.A.; and Danaher, B.G. Modifying smoking behavior of teenagers: A school-based intervention. *Am J Public Health* 70(7):722-725, 1980.

Pipher, J.R., and Rivers, C. The differential effects of alcohol education on junior high school students. *J Alcohol Drug Educ* 27:73-88, 1982.

Robinson, J. A comparison of three alcohol instruction programs on the knowledge, attitudes, and drinking behaviors of college students. *J Drug Educ* 11(2):157-166, 1981.

Rozelle, G.R. Experimental and cognitive small group approaches to alcohol education for college students. *J Alcohol Drug Educ* 26:40-54, 1980.

Sadler, O.W., and Dillard, N.R. A description and evaluation of TRENDS: A substance abuse education program for sixth graders. *J Educ Res* 71:171-175, 1978.

Schaps, E.; Moskowitz, J.M.; Malvin, J.H.; and Schaeffer, G.A. Evaluation of seven school-based prevention programs: A final report on the NAPA Project. *Int J Addict*, in press.

Sohn, M.F. Change in factual knowledge and reported use of illicit drugs resulting from the viewing of a motion picture. *Dissertation Abstr Int* 37:790A, 1976.

Stuart, R.B. Teaching facts about drugs: Pushing or prevention? *J Educ Psychol* 66(2):180-201, 1974.

Toennies, J.E. Effectiveness of selected treatment in a drug education program for university freshmen. *Dissertation Abstr Int* 33:288B, 1971.

Weaver, S.C., and Tennant, F.S. Effectiveness of drug education programs for secondary school students. *Am J Psychiatry* 130(7):812-814, 1973.

Williams, A.F.; DiCicco, L.M.; and Unterberger, H. Philosophy and evaluation of an alcohol education program. *Q J Stud Alcohol* 29:685-702, 1968.

AUTHORS

Anthony Biglan, Ph.D.
Research Scientist

Donald Hood, B.A.
Research Assistant

Paul Brozovsky, M.S.
Research Analyst

Linda Ochs, M.S.
Project Coordinator

Dennis Ary, Ph.D.
Research Scientist

Carol Black, M.S.
Project Coordinator

Oregon Research Institute
1899 Willamette
Eugene, OR 97401

140148

Increasing the Validity of Self-Report Data in Effectiveness Trials

Susan G. Forman and Jean Ann Linney

INTRODUCTION

As awareness and concern about alcohol and other drug use have grown, research and intervention efforts also have increased. Efforts to reduce alcohol and other drug use in the population have been pursued on several fronts, ranging from large-scale integrated programs to isolated one-shot interventions. The vast majority of these efforts have been implemented with little or no research or evaluation on their effects. As the storehouse of prevention and intervention programs has grown, so has concern for identifying what types of efforts work, with what populations, and under what circumstances. Flay (1985) and others have called for increased rigor in the research methodology applied to the study of alcohol and other drug use prevention and health promotion efforts. There also has been considerable debate about strategies for assessing the outcomes of these programs, typically reductions in alcohol and other drug use.

The primary assessment strategy for alcohol and other drug use behaviors is the individual self-report. Because these behaviors are illegal, the veridicality and validity of these self-reports may be questionable. There is considerable evidence that adults underreport socially undesirable or unacceptable behavior (Harrell 1985) and overestimate the incidence of socially acceptable or self-enhancing behaviors. Self-reports of alcohol and other drug use from children and youth are complicated by these same factors, albeit in unpredictable ways. In some contexts youth may overreport their alcohol and other drug use, perceiving this as the socially desirable response. Some may overreport to appear uncooperative or older. In other situations youth may fear exposure or the threat of unknown consequences for their alcohol and other drug use and underreport its incidence. Population-based surveys that assess the incidence and prevalence of specific alcohol and other drug use demonstrate that when confidentiality is ensured, surveys are administered with adequate privacy, and questions are presented in a format facilitating accurate recall and minimizing

response bias, reasonable validity can be achieved with self-report methods (Johnston and O'Malley 1985).

The self-report alcohol and other drug use survey remains the most common and perhaps the most useful instrument for evaluating the effects of preventive intervention; however, in the intervention context new problems arise. When assessment is linked to participation in an intervention program, other factors may contribute to bias in self-reports. Participants may underreport alcohol and other drug use at posttest assessment because the intervention has established an expectancy for change. Perceptions of the intervention program also may bias self-reports at pretest. If potential participants think that level of alcohol and other drug use will determine selection for the program, they may overreport or underreport their level of alcohol and other drug use depending on the perceived desirability of program participation and their hypotheses about selection criteria.

Psychological processes such as observation, storage, and short- and long-term memory can influence response to self-report questionnaires in addition to the intentional distortions already mentioned. The individual's skill at self-observation and ability to store and retrieve these observations can affect responses to self-report questionnaire items. These factors can be especially salient biasing effects when the respondents are children and adolescents.

ASSESSMENT ISSUES IN EFFICACY AND EFFECTIVENESS TRIALS

Flay (1986) has drawn distinctions among several phases of research and development of health promotion programs. Specifically, he contrasts *efficacy* and *effectiveness* trials. "Efficacy trials provide tests of whether a technology, treatment, procedure, or program does more good than harm when delivered under optimum conditions. Effectiveness trials provide tests of whether a technology, treatment or procedure, intervention or program does more good than harm when delivered under real-world conditions" (Flay 1986, p. 451). This framework distinguishes between a controlled laboratory test of a program with optimal conditions supporting the program and full participation of the target subjects and the implementation of a program in a naturalistic setting with the constraints and intrusions that the setting imposes.

Flay's (1986) model of program development and evaluation prescribes a strategy of testing program effects in which the efficacy of the intervention is established in a controlled situation with optimal conditions. Within this framework, at least theoretically, the causal link between the program and reduction in alcohol and other drug use can be demonstrated before the

program is implemented on a larger scale in the real-world setting. Thus, the effectiveness trial seeks to examine the potency of the intervention to accomplish some reduction in alcohol and other drug use given the distractions, dilutions, and countervailing forces presented by the natural environment.

Prevention efforts increasingly are being directed to high-risk groups. The real-world constraints of effectiveness trials are compounded when a high-risk group is the target of intervention. Generally, those at risk are harder to engage in the intervention and more difficult to monitor over time and may have more difficulty with standard paper-and-pencil assessments. The reduced level of control in effectiveness trials increases the need for valid assessments to have confidence in the conclusions that can be drawn from those trials.

GENERAL STRATEGIES FOR INCREASING VALIDITY OF SELF-REPORT

Several general strategies for increasing the validity of self-report data have been proposed (Rouse et al. 1985; Cone and Foster 1982). These include the design of a self-report instrument with clear, unambiguous time-and-event-grounded items and the use of additional procedures to make it more difficult or more risky for subjects to misrepresent or withhold information about their behavior. Researchers can increase confidence in the validity of self-report data by including multiple items assessing the same or contingent behaviors in the survey instrument, increasing the number of data sources and informants, and by including more than one method of measurement and using measures of multiple behaviors that are correlated with the desired outcome.

Although not previously used in alcohol and other drug use prevention studies, self-monitoring techniques may be used to increase the validity of self-reports. Self-monitoring procedures have been used to increase the accuracy and specificity of self-reports from children and adults. These procedures may be useful in further enhancing the validity of self-reports of alcohol- and other drug-related behavior. Self-monitoring includes self-observation and self-recording of individual behavior (Haynes 1978). Self-monitoring procedures typically include instruction in defining, observing, and recording a target behavior and thus may lead to more accurate self-reports of alcohol and other drug use in situations where inaccuracy is a result of questionnaire characteristics or individual cognitive processes. In a typical self-monitoring procedure, the subject records instances of the target behavior for a specified duration (e.g., 1 day, 1 week). A tally mark is placed on a recording sheet, or a counter may be used to record specific, well-defined behaviors.

Self-monitoring should be taught to children and adolescents to implement it appropriately. Mahoney (1977) suggests that the following steps be used in teaching self-monitoring procedures: (1) give specific definitions and examples of the target behaviors; (2) give specific self-recording instructions; (3) illustrate self-recording on a sample form; (4) ask subjects to repeat back definitions of target behaviors and self-recording instructions; and (5) provide trial examples of situations for subjects to self-record.

Several studies have found that individuals who are trained in self-monitoring procedures produce more accurate self-reports than those who are not (Nelson et al. 1980; Shapiro et al. 1980). Although self-monitoring can produce relatively accurate assessments of behavior by children and adolescents, reactivity of the procedure may be problematic. Reactivity occurs when self-monitoring results in behavior change without the aid of additional intervention. Some studies have documented reactivity, although all investigators have not found consistent effects (Shapiro 1984).

INCREASING VALIDITY OF SELF-REPORT WITH PHYSIOLOGICAL INDICATORS

The strategies for increasing the validity of self-reports of alcohol and other drug use in efficacy trials have tended to focus on increasing the veridicality of the subjects' responses with a "bogus pipeline procedure" (Jones and Sigall 1971) or gathering physiological indices of alcohol and other drug use, typically thought of as relatively definitive measures of alcohol and other drug use and validity checks on self-report. Neither of these procedures presents a completely valid measure, and each has serious limitations for use in effectiveness trials.

Bogus Pipeline Procedures

Bogus pipeline procedures (Jones and Sigall 1971) assume that adolescents will be more accurate in their self-reports if they believe that an objective measure of the target behavior, attitude, or belief also is being used. Murray and colleagues (1987) contend that the bogus pipeline offers the best assurance of a valid assessment of smoking, although some studies have not produced significant enhancement of self-report using this procedure. Murray and colleagues (1987) emphasize two conditions as necessary for effectiveness of the bogus pipeline: The behavior must be socially undesirable, and the subjects must believe that the investigator has a valid method to assess the behavior.

Welch and coworkers (1987) examined the effects of a bogus pipeline procedure on self-reports of alcohol and other drug use in addition to tobacco use. They concluded that the bogus pipeline is most likely to be useful in increasing adolescent self-reports of tobacco and possibly alcohol use when social norms are sufficiently strong to encourage underreporting, but that broad application could not be recommended based on the results of their study and others in the literature. They further contend that bogus pipeline methods cannot address shortcomings of self-report such as determining low-frequency use, errors in recalling quantity and frequency, and inability to determine alcohol and other drug use topography such as "sip" or "puff" rates.

Physiological Measures

Three types of physiological measures have been used in smoking prevention research: expired air carbon monoxide (CO), saliva thiocyanate, and cotinine. These have been found to have moderate correlations with self-reported smoking (Pechacek et al. 1984a). However, these measures involve several problems that decrease validity and/or make them inappropriate for use in effectiveness trials.

First, consumption of leafy vegetables produces substantial increases in saliva thiocyanate leading to false positives (Pechacek et al. 1984b). Thiocyanate samples have been found to deteriorate if they are not stored in airtight containers (Prue et al. 1981). Expired air CO has been found to have a half-life as short as 4 hours (Benowitz 1982), making it difficult to use in studies with large samples. Cotinine has greater specificity than thiocyanate and a longer half-life than expired air CO; however, the high cost of this test makes its use prohibitive in large-scale effectiveness trials. Furthermore, none of these methods is accurate in assessing low-rate smoking (Biglan and Ary 1985). Thus, there is limited utility for primary prevention efforts and intervention with younger age groups beginning experimentation.

A second issue with regard to the use of physiological measures in effectiveness trials is the effect on participation rates. Severson and Ary (1983) reported that they were required by a school district to obtain active consent from parents, rather than the more easily obtainable passive consent, only for assessments of smoking that included collection of expired air CO and saliva thiocyanate. Students whose parents consented were significantly less likely to self-report cigarette and marijuana smoking than those whose parents did not consent. Thus, individuals at greatest risk may be less likely to participate in an intervention when active consent procedures are required because of the use of physiological measures. Similarly, use of urinalysis to verify self-reports of

alcohol and other drug use probably would reduce participation. Because of the controversial nature of these tests, their inclusion in an assessment protocol might deter school districts from participation in an effectiveness study. Use of physiological measures do not appear to be an appropriate means of improving the validity of self-report in large-scale effectiveness studies. Inclusion of these procedures may introduce other threats to the validity of effectiveness trials by limiting participation.

CONSTRUCT VALIDITY MODELS FOR EFFECTIVENESS TRIALS

Once intervention programs move into the natural setting, measurement strategies intended to focus microscopically on actual rates of alcohol and other drug use not only may be impractical but also less desirable because they introduce additional unintended threats to validity. A general strategy of assessment to enhance validity can be derived from the measurement models of construct validity (Cronbach and Meehl 1955; Campbell and Fiske 1959) in which the validity of a given instrument is established by examining its covariation with other variables or instruments that theoretically predict relationships with the target variable. This model involves multiple measures, multiple informants, and assessment of secondary indices correlated with the target variable. This strategy in effectiveness trials would include informants such as peers, parents, or teachers; additional methods such as direct observation; and collection of archival data on related indicators.

Peer Ratings

Sociometric procedures have been used in numerous studies to obtain data on peer popularity, friendship, social adjustment, and social competence in children. Peers have been found to be relatively accurate, reliable raters of behavior (Hops and Lewin 1984). Although not commonly incorporated in, adolescent alcohol and other drug use research, peer ratings may be a source of potentially useful data on adolescent positive and negative social behaviors that have been shown to be correlated with that use. Whereas parents and teachers frequently are unaware of adolescent alcohol and other drug use and typically are not observers of incidents of use, peers frequently are observers of such occasions. Thus, they may provide a potential source of information concerning levels of alcohol and other drug use in the population. Peer rating procedures usually involve providing each student in a class with a list of their classmates and asking them to rate each individual on the list on a specified behavior or characteristic or, alternatively, asking students to nominate a peer who best fits some descriptor (e.g., the person you would like to work with on a class project). These procedures could be used in effectiveness trials of alcohol

and other drug use prevention programs by asking peers to rate classmates' level of social skills, peer resistance skills, or assertiveness. Descriptors specific to alcohol and other drug use situations could be included such as, "How likely is X to try to avoid drinking alcohol at a party?" These data would be examined for change in level of skill or other characteristics known to covary with alcohol and other drug use.

There are a variety of ethical concerns with the use of peer ratings of negative characteristics that must be considered carefully when this method of assessment is used in effectiveness research. Asking students to rate classmates on negative characteristics may promote future negative interactions (Asher and Hymel 1981) and contribute to negative labeling. Thus, use of negative characteristics in peer rating scales should be avoided whenever possible. However, a variety of positive social behaviors of relevance to alcohol and other drug use prevention programs (e.g., assertiveness, social skills) and other correlates of nonuse (e.g., positive attitudes toward school, involvement in school activities) can be included in peer rating scales.

Parent and Teacher Ratings

Behavior ratings completed by parents and teachers provide another source of information concerning child and adolescent behavior relevant to alcohol and other drug use prevention programs. It generally is felt that adult ratings provide a degree of objectivity that may be lacking in self-reports of children and adolescents. Parents and teachers can be appropriate sources of information because they spend large amounts of time with children and adolescents. Parents have the opportunity to observe their children in a variety of settings and situations. Teachers observe students in a relatively standard environment that allows them to make comparisons among age-related peers. Although adults are not exposed to adolescent alcohol and other drug use behaviors, they are exposed to a variety of other child and adolescent social behaviors that are related to alcohol and other drug use such as aggression and social withdrawal. As such, these behaviors are of interest in the evaluation of effectiveness of alcohol and other drug use prevention programs.

There are hundreds of behavior rating scales for children and adolescents. The Child Behavior Checklist (Achenbach and Edelbrock 1979) and the Behavior Problem Checklist (Quay and Peterson 1987) are among the most widely used and well developed. Some scales assess a variety of behavioral problems; others assess a single problem area such as conduct (Eyberg 1980) or self-control (Kendall and Wilcox 1979).

As with self-reports, parent and teacher ratings may be affected by factors such as social desirability or reactivity. However, even if parent or teacher ratings do not reflect the child's actual behavior, they can be viewed as a means of assessing social validity in that they can be indicators of whether the parent or teacher still sees the child as deviant after the intervention or perceives the intervention to have been helpful.

Community surveys of citizen perceptions of alcohol and other drug use in their neighborhoods or among youth with whom they are familiar may be another assessment strategy with social validity. Community members can be sensitive indicators of alcohol and other drug use outside of the school setting. Their direct observations of the behavior of neighborhood youth and their perception of the level of alcohol and other drug use formed by these observations and by conversations with other community members can further validate data from other sources and constitute another informant source.

Direct Observations of Behavior

Direct observation of behavior has been regarded as the "ultimate validity criterion" (Wildman and Erickson 1977). Although it is highly unlikely that observations can be used to validate actual alcohol and other drug use, this method can be used to obtain valid measurement of alcohol and other drug use-related social behaviors such as aggressiveness, withdrawal, assertiveness, and peer resistance.

Frequency recording, duration recording, or interval recording can be conducted in the natural environment, in an analog setting, or in a laboratory setting. Observation in analog settings has been particularly useful in assessing social skills learned in intervention programs. Direct observations can be conducted while the behavior is occurring or from audiotapes or videotapes. Several issues must be considered if direct observation procedures are to be conducted appropriately to yield reliable and valid results. Observer training and observer monitoring are necessary to avoid observer bias and drift. Barton and Ascione (1984) suggest a three-step training process: (1) learning the operational definitions and the recording system, (2) demonstrating mastery of the system using prerecorded tapes of behavior, and (3) demonstrating mastery of the system with *in vivo* observations. Observer training and monitoring is time-consuming and entails additional expense; however, valid and reliable data are unlikely if these procedures are not implemented. Of all assessment methods presented in this chapter, direct observation is the most personnel intensive and therefore is probably the most expensive to implement. Therefore, in

large-scale effectiveness studies it may be necessary to use direct observation for only a subsample of subjects to keep costs reasonable.

Archival Data and Secondary Indices

If alcohol and other drug use in the target population has changed, there is likely to be parallel change in other indices directly related to that use. These secondary indicators include the number of alcohol-related traffic crashes in the age group targeted, the number of alcohol and other drug incidents on school grounds, the number of referrals to school counselors or other agencies for alcohol and other drug problems, the number of other drug- and alcohol-related arrests involving the targeted age group, and the number of underage drinking arrests. These indicators would not be considered assessment of change at the individual level of program effect, but to the extent that level and intensity of alcohol and other drug use change in the population, these secondary indicators provide evidence of some degree of change at the population or community level.

Each type of indicator is available in archival data sources and can be disaggregated to examine relative change in the group targeted compared with other groups. Comparative analysis also could examine differential change in variables expected to be affected by an intervention and variables that should not be affected. This strategy of program effectiveness analysis is described in detail by Cook and Campbell (1979) and illustrated by Ross and colleagues' (1970) evaluation of the British breathalyser crackdown on alcohol-related traffic crashes. Large population-based alcohol and other drug use prevention efforts could be examined with a similar strategy.

Alcohol and other drug use prevention and health promotion program effectiveness also should be evidenced in indicators such as rates of sc ol disciplinary incidents and referrals to school counseling services. These indicators are not as directly related to alcohol and other drug use as, for example, alcohol-involved traffic crashes, but have been shown consistently to be correlated with that use at the individual level.

CONCLUSIONS

When prevention programs are tested in field-based effectiveness trials, the measurement of outcomes and effects presents some unique validity concerns. The common strategy of enhancing validity by including some physiological measure may not be feasible in effectiveness studies. Furthermore, the inclusion of physiological indicators and procedures may introduce alternative

threats to the validity of the research designs by nonrandomly affecting participation rates and reducing the overall sample size. An alternative direction to enhance the validity of assessment in effectiveness trials is the inclusion of multiple measures, multiple methods of assessment, and multiple informants or sources of data. Such a construct validity approach should be included in efficacy trials as well to establish the construct validity of each of the indices so that inferences about effects may be drawn with more confidence.

The evidence for the validity of self-report surveys should not be underestimated. There are significant advantages to the use of self-report surveys in effectiveness trials. When the conditions of assessment provide adequate confidentiality and privacy and the items of the survey are constructed to be clear and time and event bound, a self-report survey instrument can have satisfactory validity. The validity of these self-reports may be enhanced further by providing an appropriate orientation to the respondents and some prior attention to self-monitoring and self-observation procedures. Inclusion of additional measures and measurement strategies can enhance further the construct validity of the survey.

Issues of validity cannot be solved exclusively by improvements in measurement. Threats to validity related to expectancies, reactivity, and veridicality can be addressed with an appropriate study design, including no-treatment control groups and placebo control groups, in addition to replications of the intervention. The combination of appropriate designs and psychometrically sound assessment instruments is essential to validity.

REFERENCES

Achenbach, T.M., and Edelbrock, C.S. The child behavior profile: II. Boys 12-16 and girls aged 6-11 and 12-16. *J Consult Clin Psychol* 47:223-233, 1979.

Asher, S.R., and Hymel, S. Children's social competence in peer relations: Sociometric and behavioral assessment. In: Wine, J.D., and Smye, M.D., eds. *Social Competence.* New York: Guilford Press, 1981.

Barton, E.J., and Ascione, F.R. Direct observation. In: Ollendick, T.H., and Hersen, M., eds. *Child Behavior Assessment.* New York: Pergamon Press, 1984.

Benowitz, N.L. "Biochemical Measures of Tobacco Smoke Consumption." Paper presented at working group meeting on Measurement Issues in Cigarette Smoking Research, Bethesda, MD, 1982.

Biglan, A., and Ary, D.V. Methodological issues in research on smoking prevention. In: Bell, C.S., and Battjes, R., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 170-195.

Campbell, D.T., and Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 56:81-105, 1959.

Cone, J.D., and Foster, S.L. Direct observation in clinical psychology. In: Butcher, J.N., and Kendall, P.C., eds. *Handbook of Research Methods in Clinical Psychology.* New York: John Wiley & Sons, 1982.

Cook, T.D., and Campbell, D.T. *Quasi-Experimentation. Design and Analysis Issues for Field Settings.* Chicago: Rand McNally Publishing, 1979.

Cronbach, L.J., and Meehl, P. Construct validity in psychological tests. *Psychol Bull* 52:281-302, 1955.

Eyberg, S.M. Eyberg child behavior inventory. *J Clin Child Psychol* 9:29, 1980.

Flay, B.R. What we know about the social influences approach to smoking prevention: Review and recommendations. In: Bell, C.S., and Battjes, R., eds. *Prevention Research: Deterring Drug Abuse Among Children and Adolescents.* National Institute on Drug Abuse Research Monograph 63. DHHS Pub. No. (ADM)87-1334. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 67-112.

Flay, B.R. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med* 15:451-474, 1986.

Harrell, A.V. Validation of self-report: The research record. In: Rouse, B.A.; Kozel, N.J.; and Richards, L.G., eds. *Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity.* National Institute on Drug Abuse Research Monograph 57. DHHS Pub. No. (ADM)88-1402. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 12-21.

Haynes, S.N. *Principles of Behavioral Assessment.* New York: Gardner Press, 1978.

Hops, H., and Lewin, L. Peer sociometric forms. In: Ollendick, T.H., and Hersen, M., eds. *Child Behavior Assessment.* New York: Pergamon Press, 1984.

Johnston, L.D., and O'Malley, P.M. Issues of validity and population coverage in student surveys of drug use. In: Rouse, B.A.; Kozel, N.J.; and Richards, L.G., eds. *Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity.* National Institute on Drug Abuse Research Monograph 57. DHHS Pub. No. (ADM)88-1402. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985. pp. 31-54.

Jones, E.E., and Sigall, H. The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychol Bull* 76:349-364, 1971.

Kendall, P.C., and Wilcox, Z.E. Self-control in children: Development of a rating scale. *J Consult Clin Psychol* 47:1020-1029, 1979.

Mahoney, M.J. Some applied issues in self-monitoring. In: Cone, J.D., and Hawkins, R.P., eds. *Behavioral Assessment: New Directions in Clinical Psychology.* New York: Brunner/Mazel, 1977.

Murray, D.M.; O'Connell, C.M.; Schmid, L.A.; and Perry, C.L. The validity of smoking self-reports by adolescents: A reexamination of the bogus pipeline procedure. *Addict Behav* 12:7-15, 1987.

Nelson, R.D.; Hay, L.R.; Devany, J.; and Koslow-Green, L. The reactivity and accuracy of children's self-monitoring: Three experiments. *Child Behav Ther* 2:1-24, 1980.

Pechacek, T.F.; Fox, B.; Murray, D.M.; and Leupker, R.V. Review of techniques for measurement of smoking. In: Matarazarro, J.D.; Weiss, S.M.; Herd, J.A.; and Miller, N.E., eds. *Behavioral Health: A Handbook of Health Enhancement and Disease Prevention.* New York: John Wiley & Sons, 1984a.

Pechacek, T.F.; Murray, D.M.; Leupker, R.V.; Mittlemark, M.B.; Johnson, C.A.; and Schulz, J.M. Measurement of adolescent smoking behavior: Rationale and methods. *J Behav Med* 7:123-140, 1984b.

Prue, D.M.; Martin, J.E.; Hume, A.S.; and Davis, N.S. The reliability of thiocyanate measurement of smoking exposure. *Addict Behav* 6:99-105, 1981.

Quay, H.C., and Peterson, D.R. *Manual for the Behavior Problem Checklist.* Coral Gables, FL: University of Miami, 1987.

Ross, H.L.; Campbell, D.T.; and Glass, G.V. Determining the social effects of a legal reform: The British "breathalyser" crackdown of 1967. *Am Behav Sci* 13:493-509, 1970.

Rouse, B.A.; Kozel, N.J.; and Richards, L.G., eds. *Self-Report Methods of Estimating Drug Use: Meeting Current Challenges to Validity.* National Institute on Drug Abuse Research Monograph 57. DHHS Pub. No. (ADM)88-1402. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1985.

Severson, H.H., and Ary, D. Sampling bias due to consent procedures with adolescents. *Addict Behav* 3:433-437, 1983.

Shapiro, E.S. Self-monitoring procedures. In: Ollendick, T.H., and Hersen, M., eds. *Child Behavior Assessment.* New York: Pergamon Press, 1984.

Shapiro, E.S.; McGonigle, J.J.; and Ollendick, T.H. An analysis of self-assessment and self-reinforcement in a self-managed token economy with mentally retarded children. *App Res Ment Retard* 1:223-240, 1980.

Welch, C.E.; Gorman, D.R.; Marty, P.J.; Forbess, J.; and Brown, B. Effects of the bogus pipeline on enhancing validity of self-reported adolescent drug use measures. *J School Health* 57:232-236, 1987.

Wildman, B.G., and Erickson, M.T. Methodological problems in behavioral observations. In: Cone, J.D., and Hawkins, R.P., eds. *Behavioral Assessment: New Directions in Clinical Psychology.* New York: Brunner/ Mazel, 1977.

AUTHORS

Susan G. Forman, Ph.D.
Associate Provost
Office of the Provost
University of South Carolina
Columbia, SC 29204

Jean Ann Linney, Ph.D.
Associate Professor
Department of Psychology
University of South Carolina
Columbia, SC 29204

1401H9

Technology Transfer

Steven P. Schinke and Mario A. Orlandi

INTRODUCTION

Before 1865, scurvy was the leading cause of death among sailors, exemplified by Vasco de Gama's voyage around the Cape of Good Hope in 1497. Of the crew of 160 sailors, 2 of 3 died of scurvy during the trip. Due to the frequency of such fatalities, James Lancaster, an English sea captain, carried out a field experiment in 1601 on the efficacy of lemon juice to prevent scurvy (Rogers 1983).

Though hardly a randomized clinical trial, the field study that Lancaster designed allowed him to compare the incidence of scurvy among sailors who received a daily ration of lemon juice with sailors in a control group who received no lemon juice. At the end of the trial, all the sailors who took lemon juice daily remained healthy; of the 278 sailors who were not given lemon juice, 110 (40 percent) died from scurvy. Despite these impressive results, Lancaster's study did not lead to systematic use of citrus products to prevent or treat scurvy in the British Navy. Scurvy, in fact, went untreated and neglected as a topic for further intervention efforts for the next 150 years.

In 1747, a British Navy physician, James Lind, who knew of Lancaster's findings, conducted another experiment to evaluate citrus products as a treatment for scurvy. Lind studied five diets for treating scurvy among sailors aboard the HMS Salisbury. Those diets consisted of daily doses of (1) two oranges and a lemon, (2) a half-pint of sea water, (3) six spoonfuls of vinegar, (4) a quart of cider, or (5) nutmeg or vitriol elixir. Except for the sea water, which served as a placebo control, each of the other diets had potential merit for curing scurvy. However, only the sailors who received the citrus fruits were cured (Rogers 1983).

As with Lancaster's study, the results of Lind's experiment did not usher in a widely accepted and adopted preventive and curative intervention for scurvy among sailors. Not until 48 years after Lind's study did the British Navy require sailors to receive a daily ration of citrus. The British Board of Trade waited another 70 years to adopt a similar policy and, thus, completely eradicated

scurvy among all British sailors. In sum, a proven technology for preventing scurvy in British naval operations took 264 years to move from initial field research results to wide-scale implementation (table 1).

TABLE 1. *Milestones in the prevention of scurvy among British sailors*

Date:	1601	1747	1795	1865
Event:	Lancaster study	Lind study	Navy adopted	Merchant adopted

What lessons can we learn from the failed technology transfer evident in the British Navy's adoption of a known preventive intervention to eradicate scurvy among crews at sea? More important, how can we apply those lessons to ensure the transfer of preventive intervention technologies in the field of substance abuse?

This chapter defines technology transfer and describes and illustrates the stages of technology transfer in the field of substance abuse prevention. Next, the authors note the relevance of concepts surrounding (1) innovation and change and (2) diffusion for transferring preventive intervention technologies. The chapter next considers factors that influence technology transfer. Finally, the authors present and discuss strategies to enhance the technology transfer of interventions for preventing substance abuse.

TECHNOLOGY TRANSFER

Definitions

Because vagaries in terminology have the potential to confuse and contradict, a brief review of the meanings and purposes of technology transfer seems in order. Technology transfer, as employed in this chapter, represents a process through which methodologies and interventions for substance abuse prevention are born of research and move into application. Through that process, ideas are generated, innovations are created then tested, and results are disseminated to and, ideally, adopted by practitioners and service providers.

Definitions of technology transfer range from the general to the specific. Dans (1977) offered a general definition when he stated that technology transfer is "shorthand for the diffusion of technology from its discovery to its appropriate application." A specific definition is illustrated by Brown and colleagues (1979),

who define technology transfer as "instances where the given technology moves from one situation to another, which may require changes in the technology, the content to which it is moved, or both." The technology transfer process lends itself to the development and diffusion of any innovation, including interventions, measurement instruments, assessment procedures, taxonomies, and treatment guidelines.

Stages of Process

Based on the aforementioned literature and research and drawing on our colleagues' and our own original data on the development and application of preventive interventions for substance abuse problems, technology transfer is conceptualized as occurring in eight sequenced stages (table 2).

TABLE 2. *Stages of technology transfer and drug abuse preventive intervention*

Technology Transfer Stages	Drug Abuse Prevention Stages
1. Basic Research	Theory development; data synthesis
2. Applied Research	Case study; clinical work; exploratory studies
3. Technology Development	Construction of intervention curriculums
4. Evaluation	Clinical trials; analog and outcome
5. Demonstration	Field studies with evaluation; focus on population
6. Adoption	Use in nonresearch settings
7. Application in Practice	Widespread acceptance and use of intervention among line practitioners
8. Obsolescence	Disuse of old intervention technology; evolution to a new technology

SOURCE: Adapted from Office of Technology Assessment, 1982.

Uncommonly associated with substance abuse prevention efforts, *basic research* is no less essential to the creation, design, and synthesis of new technologies. Basic research is the stage when investigators discover relationships and make links between one body of knowledge and another to

suggest directions for preventive intervention development and research. Examples of relationships and links are initial work on psychological "inoculation" theory, problem behavior syndrome, and social learning theory. Each of these areas of research has resulted in milestones for understanding and preventing substance abuse behavior.

In the field of substance abuse prevention, *applied research* is arguably the stage of theoretical work, clinical use, or case study at which innovations show their greatest promise. Applied research with clinical samples is the time when investigators receive feedback on the relevance and potential efficacy of innovations and responsively adapt interventions based on that feedback. Examples of applied research are plentiful in the literature and are illustrated by investigations of social and interpersonal skills as a way of understanding peer pressure toward and away from substance use and by investigations of procedures for assessing and enhancing the accuracy of adolescents' self-reported substance use behavior.

Within the field of substance abuse prevention research, *technology development*, the third stage of transfer, often focuses on the crafting of theoretically sound and clinically viable intervention curriculums. Important research at this stage usually benefits from external funding and from a series of investigations conducted by the same or, more often, different investigators. An example is the development of skills-based technologies for preventing substance use among adolescents. At several laboratories across the country, investigators concurrently and sequentially develop and refine skills approaches to preventive intervention among youth at risk for tobacco, alcohol, and other drug use.

Evaluation is the stage of technology transfer that occupies the bulk of resources and time for substance abuse prevention investigators. Here, investigators conduct controlled analog studies or take their interventions into applied settings for controlled outcome research. Evaluations of preventive interventions in the field of substance abuse invariably include randomized designs, careful measurements, and comparison or control interventions. Increasingly, the randomized clinical trial is the accepted design for outcome evaluation research on the efficacy of interventions for preventing substance abuse.

Demonstration, the fifth stage of technology transfer, moves beyond the clinical trial of the evaluation stage and includes the widespread application of a prevention technology within a research design. Although definitions for demonstration research vary depending on the purpose of the study, it is often part of the intervention outcome activity. Demonstration projects also are

demarcated by a focus on a target population for whom the intervention was intended. Examples of demonstration studies are projects sponsored by the Office for Substance Abuse Prevention aimed at high-risk youth and allotting a relatively small fraction of resources to evaluation, relative to costs allocated to program implementation.

Adoption, the sixth stage in the technology transfer process, begins the final dimension. Recognizing the novelty of concepts surrounding this stage and the two stages to follow, we devote detailed discussion in this section to adoption, application in practice, and obsolescence as they relate to substance abuse preventive interventions. To set the stage for that discussion, brief mention of the salience of adoption is warranted. The *raison d'être* of substance abuse prevention research is to develop and empirically test interventions that will help people altogether avoid problems with alcohol and other drugs and other harmful substances. If those interventions are not adopted in nonresearch settings, much of the work in their development and testing has been ill spent. Thus, adoption for our purposes represents that stage at which interventions are employed by constituencies other than investigators in the service of preventing substance abuse.

Because adoption is a necessary *application in practice*, this seventh stage of the technology transfer process draws attention to the need for diffusion of innovative ideas, techniques, and strategies. As with the stage of adoption, more remains to be said about application in practice later in this chapter. For now, we use an example to differentiate adoption from application in practice. Our current research is concerned with the crafting and evaluation of interventions to prevent drug use among Native American adolescents and is occurring exclusively in the Pacific Northwest.

We expect our curriculum to have passed major field tests within the next couple of years at which point it will likely enjoy modest adoption in the Northwest. Our confidence in the adoption of the curriculum derives from the intimate knowledge we possess of most tribal groups in the Northwest, the reception to date that the curriculum has received, and the results of efforts to train many Native professionals in the delivery of the intervention. Yet, our curriculum will not move into the application stage until we or another interested party expressly plan for and undertake the diffusion of the curriculum throughout Indian communities that will benefit from it most. Those steps depend on our ability to produce an inexpensive product and then to market, advertise, and distribute it successfully.

The final stage of the technology transfer process is *obsolescence*, which because of its inevitability also must command the attention of investigators

who design and test preventive interventions for substance abuse. Simply stated, obsolescence means that a technological innovation is no longer used. To connote obsolescence as simply discontinued use of the innovation, however, may obscure the importance of this final stage in the generation of new technologies to address new problems or improve on solutions to existing problems. The discussion that follows addresses the benefits of the obsolescence stage by relating it to the importance of the evolution from one technology transfer process to another.

INNOVATION AND CHANGE

A set of concepts integrally related to technology transfer concerns the implementation of innovation and change, which are concepts imbedded in stages of adoption and application in practice. In a seminal paper on the implementation process through which innovation and change occur, Kolbe and Iverson (1981) identified five phases of programmatic implementation for social and health education efforts. These phases are mobilization, adoption, implementation, maintenance, and evolution.

Mobilization defines the time when service providers consider the option of using a new program or improving an existing one. To become mobilized, for example, school administrators or community agency program planners and clinicians might recognize deficits in their current efforts to combat and prevent drug use. Alternatively, mobilization may occur when administrators and clinicians note a growing or new problem with drug use among their students and clients. The current climate of helplessness that surrounds problems in the use of crack cocaine in many regions of the country illustrates this latter type of mobilization influence.

Adoption is the stage during which program planners demonstrate a commitment to a new program. Relative to drug abuse prevention efforts, this stage is illustrated by administrators' and clinicians' declared acceptance of risk factors for substance use among youth, expressed belief in the modifiability of those risk factors, and stated value of preventive intervention efforts.

Implementation, in the framework laid down by Kolbe and Iverson, is defined as the time when the course of action is put into practice. Quite simply, implementation then represents the start of the program *per se*. In substance abuse prevention, program implementation covers the period of intervention delivery.

Maintenance defines the stage during which an innovative program is continued by the host organization, school, or agency. The maintenance of a drug abuse

prevention effort, for instance, occurs when the intervention program persists beyond its original implementation by virtue of a supportive administrative or clinical staff. Maintenance is clearly an important stage of innovation and thus warrants further consideration.

In their paper, Kolbe and Iverson review research in support of four factors that influence the maintenance of innovations. The first of these four factors is the degree of involvement of program staff in the intervention program, as manifested by such activities as staff participation in program implementation, curriculum development, and inservice training. The second factor is the degree of cooperation among staff members in the execution of the innovation or program. Third, new programs are maintained to the extent that staff members have available assistance for training and implementation. Fourth, maintenance of a program is associated with the level of communication among program staff.

Evolution, as a stage in innovative program implementation, occurs when the host organization changes the new intervention or practice. For instance, school officials or community agency staff members would move into the evolution stage if they broadened a program to include new objectives and curriculums. Illustrative of this is an effort to expand a drug abuse prevention curriculum to encompass health promotion content and goals.

Together, the five phases of program implementation provide a structure for anticipating and monitoring the manner in which a preventive intervention effort will be embraced, applied, and continued by school personnel and/or other human services staff. Warranting note is that the six phases of implementation identified by Kolbe and Iverson enjoy parallels with other conceptual presentations of innovation and diffusion. For example, Rogers (1983) details five stages in the innovation-decision process: knowledge, persuasion, decision, implementation, and confirmation. Because Kolbe and Iverson's model adds an important sixth stage of evolution, it appears better suited than the model of Rogers and others for the transfer of substance abuse prevention technology.

The following quote from Kolbe and Iverson concludes our coverage of this area:

> The effectiveness of health education is ultimately determined by whether it is implemented, and how it is implemented. Although a given health education innovation may be designed and experimentally assessed to promote well-being with some measure

> of effectiveness and efficiency, the actual impact of the innovation will depend upon the manner in which it is disseminated, initiated, and maintained (Kolbe and Iverson 1981, p. 78).

The significance of implementing programmatic innovations for substance abuse preventive interventions is clear in Kolbe and Iverson's conclusions about impact. Also important for substance abuse prevention research and programing are the concepts of diffusion.

DIFFUSION

Earlier in our description of the eight stages of technology transfer, we observed that the critical event of moving an innovation from adoption to application in practice was largely dependent on diffusion efforts. Therefore, the area of diffusion is key to transferring prevention technologies into the hands of those who want and need them.

Definition

According to Basch and colleagues, "*diffusion* is generically defined as the spread of new knowledge. . . . The classical formulations of diffusion occur when knowledge is seen as being generated in and emanating from a single source, moving from those who have it to those who do not" (Basch et al. 1986, p. 2). Basch and colleagues also describe the role of diffusion systems in the spread of innovations. Defined as sets of relationships among human beings and social organizations that foster the sharing of knowledge and products, diffusion systems are the vehicles for moving new ideas and practices among people, institutions, and service providers.

Active and Passive Systems

An effective diffusion system facilitates the transfer of knowledge among organizations and people in need of the knowledge through a passive or an active process. Illustrative of an active process for diffusion are outreach efforts by government and private bodies that seek to inform practitioners, service providers, and relevant institutions about innovations. When advertised and distributed nationally, this technical review of preventive intervention issues and knowledge represents an active diffusion system.

A passive system for diffusion requires the user to request information on the innovative idea, process, or practice. Examples of passive systems are online retrieval services from the Government, such as those available at the National

Library of Medicine, or from the commercial sector. Possibly, the ideal diffusion system would combine these two types of systems, exposing professionals and scientists to information sources of innovations, then allowing consumers to retrieve information in areas of greatest interest to them.

FACTORS THAT INFLUENCE TECHNOLOGY TRANSFER

Whether each of the eight stages for technology transfer occurs and the rate of movement from one stage to the next depend on several factors encompassing technology development, the target population for preventive intervention, and environmental receptivity aspects of organizations that will apply the technology. The following sections summarize extant knowledge about each of these three types of factors that influence technology transfer in substance abuse prevention intervention development, adoption, and application.

Technology Development

The range of options open to investigators of substance abuse prevention interventions is large. Besides the obvious choices of setting, target population, and intervention type, investigators must choose the nature and number of substances to include in an intervention program. Each decision made at the onset of an intervention development necessarily precludes the selection of other options that later affect the technology transfer process. For example, a program that is initially aimed at tobacco use among middle-class youth is unlikely to find a receptive audience among inner-city school administrators faced with problems of crack use among lower socioeconomic status youth.

Parallel considerations are necessary throughout the technology transfer process as investigators evaluate their interventions and subject prevention strategies to demonstration tests in the field. Unless a preventive intervention is designed for an at-risk population and a high-priority substance abuse problem, the intervention will not readily lend itself to such applications later in the transfer process. In recognition of the critical and sometimes irreparable decisions reached during the creation and development of a technology, a working group of the American Public Health Association (APHA) recently drafted and made available guidelines for prevention program design (American Public Health Association 1987). Those guidelines restate programmatic considerations discussed in much of the prevention literature and thus will not appear new to most readers. Even so, the five APHA guidelines deserve brief paraphrasing here because of their salience for substance abuse prevention intervention in enhancing technology transfer. The five guidelines are as follows:

1. Prevention programs should address one or more risk factors that are carefully defined, measurable, modifiable, and prevalent among the members of a chosen target group. The risk factors should constitute a threat to the health status and to the quality of life of target group members.

2. Prevention programs should reflect the special needs, characteristics, and preferences of target groups.

3. Programs should include interventions that will effectively reduce a risk behavior and that are appropriate for a particular setting.

4. Prevention programs should identify and implement interventions that make optimum use of available resources.

5. Prevention and health promotion programs should be organized, planned, and implemented so that their operation and effects can be evaluated.

By considering and addressing these guidelines during the creation, development, evaluation, and demonstration testing of preventive interventions, investigators in the substance abuse field will increase the likelihood of transferring their innovations and technologies. Closely aligned with considerations about the development of innovations for prevention programing are factors that concern the receipt of intervention among members of the target population.

Target Population Factors

For purposes of technology transfer, factors relevant to the target population concern the manner in which preventive intervention content is perceived by members of target groups. Research on the receipt of intervention includes work on variables in the communication of problem prevention and health promotion content. That research demonstrates that effective behavior change efforts have messages that are clear, coherent, consistent, and compatible with the values of the target group (Bloom 1987; Durlak and Jason 1984; Gullotta 1987; Orlandi 1986). From that same literature come conclusions that successful preventive interventions are sufficient to influence individuals all along the behavior change continuum.

A useful model for understanding and addressing target population factors in the transfer of prevention technologies is provided by Farquahar and colleagues (1981), whose model interventions are conceptualized as a series of communicated messages with five dimensions. These five dimensions concern the communication message, channel, source, destination, and receiver.

The *message* dimension of communication, for our purposes, includes the context, form, and structure of substance abuse prevention intervention. Effective messages in prevention programing should therefore express the language and style appropriate and culturally relevant for the target group.

Channel, as a dimension of population targeting in technology transfer, includes the characteristics of the medium or media used to convey the preventive intervention message. In substance abuse prevention programing, channels for transmitting intervention content should be those that are known to consistently reach a high percentage of the specific target group.

The *source* dimension concerns the attributes of the individual, group, or organization perceived as the origin of the preventive intervention message. Perceived sources of messages for substance abuse preventive intervention are the providers, organizations, or institutions that are respected and credible among members of the target population.

The dimension of *destination* includes characteristics of the targeted behavior change that preventive intervention seeks to instill. For instance, the destination or targeted change of substance abuse preventive intervention is often framed in terms of a reduced risk for alcohol, tobacco, or other drug use. Whatever the destination of a preventive intervention effort, it must be feasible and salient to members of the target population.

The *receiver* dimension in the current context includes pertinent attributes of the target audience. Consequently, messages in line with the implications of this dimension are constructed so that they are relevant to the cognitive ability, belief structure, and value system of the intended receivers of the substance abuse preventive intervention.

Environmental Factors

The third set of factors concerns the receptiveness of the host environment for adoption of innovative prevention programs. Addressing environmental factors for diffusion and adoption of relevant technologies, Stevens and Davis (1988) conducted a study of school districts viewed as having strong attributes in favor of health promotion curriculums, labeling them HOT districts, and those viewed as not favoring health promotion curriculums, labeling them COLD districts.

Supporting their prediction about the favorable environment toward health education in HOT school districts, Stevens and Davis found differences between the two groups of districts in their study on several dimensions. Those differences were evident from discriminant function analyses on dimensions of

staff services, staff inservice programs, and administrative behaviors. Compared with schools in COLD districts, schools in HOT districts were more apt to have positively modified traditional norms of health education, addressed social and organizational factors, placed staff development as a primary target for educational efforts, given consideration to nutrition in foods served inside and outside of the cafeteria, and extended health education services beyond the classroom.

Attempting to explain differences in school districts' receptivity to health education and health promotion efforts for students, Stevens and Davis examined educational curriculums at HOT districts and COLD districts. To their surprise, the investigators found that HOT and COLD districts had similar education programs for their students. This finding led Stevens and Davis to study further the reasons for a district's readiness for health promotion curriculums. In so doing, they learned that administrators in HOT districts, to a greater degree than their counterparts in COLD districts, reported that they preferred to use their efforts and resources to demonstrate a commitment to health concepts before investing in revisions to their curriculums.

STRATEGIES TO ENHANCE TECHNOLOGY TRANSFER

Based on preceding literature and on our original experiences, we conclude this chapter with four steps that investigators and policymakers can follow to increase the likelihood of prevention technology transfer in the substance abuse field. We call these steps stick to the basics, replicate studies, analyze costs, and strive for high-quality dissemination.

Stick to the Basics

Adapted from a similar strategy called "stick to the knitting" as advanced by *In Search of Excellence* (Peters and Waterman 1982), our initial step for transferring prevention technologies is also the most important. In this step, we recommend that prevention researchers plan and execute studies within conventional research designs of test interventions that are theory based and empirically indicated. Despite its straightforward appearance, this recommendation is not easily implemented. The tendency in the current climate for prevention research is toward the design of increasingly complex studies that attempt to outdo what has come before.

Equal pressure toward complicated designs and the exploration of new frontiers is exerted by the requirements of review groups for external funding agencies. Indeed, the likelihood appears small that a review group will act favorably on yet another controlled outcome study of an intervention to prevent drug use among

adolescents. The basis for this prediction is our own experience and our vicarious experiences with other investigators and proposals. Innovative studies with new populations and in new settings are certainly needed; yet, well-grounded, elegantly designed studies will do much for transferring prevention technology.

Replicate Studies

Long associated with the advancement of scientific knowledge, replication studies are a wise investment in technology transfer for substance abuse preventive interventions. Replications of successful interventions to prevent substance abuse serve several important functions. By replicating the interventions, methodologies, and results of their colleagues, prevention researchers can confirm the value of curriculums for substance prevention among service providers and related consumers who will eventually embrace and adopt innovative prevention programs. Replications provide ready opportunities to refine, build on, and improve interventions for substance abuse prevention.

Replication studies draw added attention to the existence and effects of preventive interventions. As scientific knowledge on tested preventions grows and becomes familiar to research and professional audiences, the likelihood of technology transfer commensurately increases. Interventions tested in many replication studies are candidates for technology transfer due to mounting evidence on their efficacy. In a fashion parallel with other scientific areas, prevention interventions for substance abuse will inexorably move into everyday practice and settings.

Analyze Costs

The wisdom of cost analyses of preventive interventions is apparent in the definition put forth by Bloom, who described such analyses as providing "research that evaluates the total benefits of some program against the total cost of some program (or some comparison program) so that decision makers can allocate limited resources to the net benefit of society" (Bloom 1986, p. 28). For present purposes, decisionmakers are those administrators, practitioners, and service providers who ultimately must use innovative prevention interventions in environments in which youth at risk for substance use reside. Cost analyses yield data to aid adoption decisions by generating what Bloom called a "common coin of exchange."

Admittedly complex to compute, cost analyses can at least indicate the price that service providers can expect to pay for a particular intervention program.

At best, cost analyses can produce a ratio of prevention intervention expenses relative to substance abuse reduction or onset delay outcomes. Guidelines for calculating intervention costs and for estimating the payoffs of intervention on program recipients and provider institutions are found in Windsor and colleagues (1984).

Strive for High-Quality Dissemination

Our last recommendation for enhancing the chances of technology transfer of substance abuse prevention interventions concerns the manner in which investigators disseminate their findings. In this recommendation, we urge investigators to publish substance abuse prevention results in the best journals and books, through popular press outlets, and via presentations at prestigious conferences. This enhancement to technology transfer will obviously bring the message of prevention research into the scientific and public eye.

Without intending to appear glib or facile, our recommendation for investigators to strive toward high-quality dissemination efforts is aimed at reminding those who create and test preventive interventions that their work is only as notable as their success in telling professionals and the lay public about their findings. The process of high-quality dissemination is not easy. Investigators must prepare for many rejections from such journals as *Science, New England Journal of Medicine, Lancet, Nature*, and *Scientific American* as they submit their prevention intervention results.

Similarly, the publication of books for both professionals and laypersons and the development and presentation of conference papers demand considerable time and attention that investigators could otherwise devote to the crafting of grant proposals and papers for the usual specialty journals. But we are confident that the payoffs of papers that appear in the best outlets will more than make up for the labor required for their production.

CONCLUSIONS

Technology transfer, for purposes of this chapter, is defined as the application of scientific knowledge from the original context—in which the knowledge, findings, or strategies were generated—to new, unresearched contexts. Within this definition, an example of technology transfer is the application of a preventive intervention strategy that has been tested and found to be successful among members of one population to a different and unstudied population. In substance abuse prevention studies, for instance, investigators often employ interventions with one population that has undergone scientific testing with another population. Such cross-population applications of

substance abuse preventive interventions raise technology transfer issues that deserve attention and demand prospective empirical research.

This chapter has enumerated and addressed several issues in the transfer of preventive intervention technologies for reducing the risks of substance use and abuse. After defining stages of the process of technology transfer, we addressed issues surrounding the implementation of innovation and change. Of the different existing models, Kolbe and Iverson's is superior for anticipating and monitoring the manner in which a prevention intervention effort will fare because other programs lack a final stage of evolution.

After discussing diffusion of and factors that influence technology transfer, we concluded this chapter with our own suggestions to enhance technology transfer. Based on extant literature and our own research, the four steps (stick to the basics, replicate studies, analyze costs, and strive for high-quality dissemination) provide a method to increase the likelihood of technology transfer in the substance abuse field.

REFERENCES

American Public Health Association. Criteria for the development of health promotion and education programs. *Am J Public Health* 77:89-92, 1987.

Basch, C.E.; Eveland, J.D.; and Portnoy, B. Diffusion systems for education and learning about health. *Fam Community Health* 9:1-26, 1986.

Bloom, M. Hygieia at the scales: Weighing the costs and effectiveness of prevention/promotion, with special reference to mental retardation. *J Primary Prev* 7:27-48, 1986.

Bloom, M. Toward a technology in primary prevention: Educational strategies and tactics. *J Primary Prev* 8:25-48, 1987.

Brown, J.; Wooten, F.T.; and Fisher, W. Technology transfer in medicine. *CRC Crit Rev Bioengineer* 4:45-70, 1979.

Dans, P.E. Issues along the Potomac: "Efficacy" and "technology transfer." *South Med J* 70:1225- 1231, 1977.

Durlak, J.A., and Jason, L.A. Preventive programs for school-aged children and adolescents. In: Roberts, M.C., and Peterson, L., eds. *Prevention of Problems in Childhood: Psychological Research and Applications.* New York: J. Wiley & Sons, 1984. pp. 103-132.

Farquahar, J.W.; Magnus, P.F.; and Maccoby, N. The role of public information and education in cigarette smoking controls. *Can J Public Health* 72(6):412-420, 1981.

Gullotta, T. Prevention's technology. *J Primary Prev* 8:4-24, 1987.

Kolbe, L.J., and Iverson, D.C. Implementing comprehensive health education: Educational innovations and social change. *Health Educ Q* 8:57-80, 1981.
Office of Technology Assessment. *Technology Transfer at the National Institutes of Health.* OTA-TM-H-10. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1982.
Orlandi, M.A. Community-based substance abuse prevention: A multicultural perspective. *J Sch Health* 56:394-401, 1986.
Peters, T.J., and Waterman, R.H. *In Search of Excellence: Lessons from America's Best-Run Companies.* New York: Warner Books, 1982.
Rogers, E.M. *The Diffusion of Innovations.* 3d ed. New York: Free Press, 1983. pp. 7-8.
Stevens, N.H., and Davis, L.G. Exemplary school health education: A new charge from HOT districts. *Health Educ Q* 15:63-70, 1988.
Windsor, R.A.; Baranowski, T.; Clark, N.; and Cutter, G. *Evaluation of Health Promotion and Education Programs.* Palo Alto, CA: Mayfield, 1984.

ACKNOWLEDGMENTS

This chapter was prepared with support from National Institute on Drug Abuse grants DA-03277 and DA-05321.

NIDA acknowledges the details and lessons of E.M. Rogers (1983) in his description of the long process of eradicating scurvy in the British Navy as a classic failure of technology transfer.

AUTHORS

Steven P. Schinke, Ph.D.
Professor
Columbia University School of Social Work
622 West 113th Street
New York, NY 10025

Mario A. Orlandi, Ph.D., M.P.H.
Division Chief
Division of Health Promotion Research
American Health Foundation
320 East 43rd Street
New York, NY 10017

Prevention Evaluation Research Methods: Findings and Consensus

Carl G. Leukefeld and William J. Bukoski

INTRODUCTION

Prevention research has been described in various ways (Leukefeld, in press). Some suggest that drug abuse prevention, and the consequent research methodology, is a scientific endeavor focused on etiology, human development, vulnerability, and evaluation research. Others indicate that drug abuse prevention research is "lightning-rod" research that has repeatedly attracted negative findings. Although this volume does not wish to enter a potential controversy, it does seek to clarify issues related to prevention evaluation methodology.

This review of drug abuse prevention research methodology reinforces the editors' belief that there is some agreement regarding the next steps in refining drug abuse prevention research methodology. The deliberations are applicable to alcohol prevention research as well as other drug abuse prevention. Finally, amid the presentations and discussions of technical issues, such as power analysis, attrition assessment, and structural equation models, two themes repeatedly surfaced as central to the continued development and application of scientifically sound prevention research methods. The first was the importance of planning experimental, quasi-experimental, and policy research to have a high degree of internal validity to increase confidence in the veracity of research results. The second theme focused on those issues, factors, and confounds in research design and implementation that threaten external validity or generalizability of research findings to a population. These two extremely important issues reflect a traditional, classical research perspective.

Clearly, drug abuse prevention evaluation has come a long way toward reaching maturity and is now striding toward adulthood. Themes reflected in this volume suggest areas for methodological fine-tuning rather than reconstructing existing research. The consensus statements presented in this chapter reflect agreement rather than diversity. Those methodological suggestions can be incorporated into existing as well as future drug abuse

prevention evaluation research. The editors hope that these suggestions will be used by investigators to refine their designs to enhance the robustness of future prevention research.

The remaining sections of this chapter present the essence of the technical review meeting. Consensus recommendations are presented at the end of the chapter. It is the editors' hope that questions asked about prevention research evaluation methodology might be answered partially by referring to this volume and that future prevention interventions might be evaluated more precisely.

DOCUMENTING THE INTERVENTION

After positing Rudner's (1966) definition of a theory as ". . . a systematically related set of statements, including some law-like generalizations, that is empirically testable," Flay and Petraitis (this volume) indicate that researchers have developed numerous theories related to drug abuse (Lettieri et al. 1980) that were derived from narrow disciplinary perspectives. Theory, in addition to specific functions, is important to discriminate between program and theoretical failure and to contribute to knowledge and research efficiency. In addition to other factors, theory is important for external validity or generalizability and construct validity or understanding immediate or delayed intervention effects. Theory also has special implications for sample size, unit of assignment, and study attrition. Nevertheless, theoretical considerations are frequently forgotten when prevention evaluation studies are planned and carried out. Clearer definitions of drug abuse prevention interventions may go a long way toward clarifying the nature of an intervention's impact. Hawkins and colleagues (1989) suggest possible goals for drug abuse prevention activities, which range from eliminating patterns of pathological use to delaying early onset. Clearly, the real or implied program goal(s) has an important relationship to outcome measures. Gilchrist (this volume) adds that defining risk for regular, frequent, and committed drug use is important but that refined diagnostic tools are not available.

However, Gilchrist suggests steps that might be useful in developing a framework of prevention goals and targets. An important consideration is matching a program's strength with an individual's vulnerability to drugs. To better focus this type of matching, three strategy levels of prevention program intensity are suggested: universal, selective, and indicated. Individual-in-environment assessment techniques can be used to specify the level of prevention interventions. For clarity, assessment and intervention targeting should be coupled with an analysis of accessing various community agencies, other than schools, that could incorporate drug abuse prevention interventions into institutional settings.

Pentz and Trebow (this volume) identify three program implementation issues that can affect the quality of prevention programs: (1) adherence to program implementation design, (2) consumer exposure to the prevention program, and (3) program change during implementation. These kinds of program implementation issues have been a problem for various drug abuse evaluation research studies and are complicated by measurement as well as environmental factors. Mark (1983) suggests that an estimate of program effects should incorporate confidence limits of program effectiveness, with program assignment at the low end and implementation at the high end of the confidence limits. Such an estimate of program effectiveness could go a long way toward incorporating implementation issues into outcome measures and should be tried in future analyses.

MEASURING THE INTERVENTION

Design choice (i.e., experimental or quasi-experimental) is clearly related to environmental factors, which frequently determine the type of prevention study. These environmental factors usually are beyond the investigator's control but can limit the possibility of making causal inferences. Taking these factors into account, Snow and Tebes (this volume) review validity issues and recommend selecting the most rigorous design that is feasible within environmental constraints. The advantages of tightly controlled experimental designs are presented by Campbell and Stanley (1963), who also provide a range of possible designs along with their related strengths and weaknesses. Clearly, design decisions should emphasize the advantages of using controlled experimental designs for feasible prevention research.

Bentler (this volume) takes another approach to rigor but recommends that it be used only after there is design breakdown or when certain experimental conditions cannot be met. For instance, when randomization cannot be used to ensure that extraneous variables are controlled, Bentler suggests structural equation modeling, which can assist researchers in overcoming design problems. If statistical assumptions are met, structural equation modeling can assist drug abuse prevention researchers to examine pretest differences and treatment effects, estimate equivalence between groups, control for missing data, and develop program participation indicators.

Dwyer and MacKinnon (this volume) consider issues related to outcome measurement and describe extrascience factors (i.e., costs, ethics, and confidentiality) as important when selecting a measurement strategy. Outcomes used in drug abuse prevention evaluations can include a variety of physical, biological, psychological, and sociological factors. Extrascience factors may drive the nature of a study and influence a study's outcome.

Outcome measurement might incorporate prevalence as well as incidence data and examine the relationships among multiple drug use outcomes, which frequently change. Thus, analyses of multiple outcomes, including biological measures, can increase the understanding of prevention intervention effects.

ASSESSING EFFECTIVENESS

The overriding goal of prevention evaluation research is to assess the effectiveness of prevention interventions and serve as the core of methodological considerations. A major issue in assessing the effectiveness of drug abuse prevention interventions is to consider the long-term effects of the intervention as well as replication. Hawkins and coworkers (this volume) outline a strategy for assessing the long-term effects of drug abuse prevention interventions in replicable studies. Agreeing with Flay and Petraitis (this volume), Hawkins and coworkers indicate that definition of theoretical constructs is critical at the outset of an intervention evaluation study. In addition, theory must be linked with the intervention(s) and the outcome(s). Assessing the degree to which the intervention is implemented frequently is overlooked and consequently weakens a study's replicability. Implementation data should be collected, analyzed using univariate as well as multivariate models, and clearly reported. It is suggested that drug abuse prevention studies should be incorporated (nested) within longitudinal panel studies. These kinds of longitudinal studies can add to knowledge about the intervention's effects and etiology.

Attrition from prevention evaluation research studies has been a routine occurrence that is frequently overlooked and unreported. Biglan and colleagues (this volume) examine other drug abuse and alcohol and smoking prevention studies for attrition rates and detail the importance of attrition on internal and external validity. They report that none of the alcohol and other drug abuse prevention evaluation studies in their review examined the differences between study dropouts and remaining subjects. In addition, only about one-fifth of smoking prevention studies included attrition analyses. On the other hand, study drop-ins also should be considered, and methods of reducing attrition should be emphasized.

A major issue in assessing effectiveness is selection of outcome measures as well as using self-report data. Forman and Linney (this volume) suggest that in field-based effectiveness trials outcome measures present unique validity issues; physiological indicators are not the *sine qua non;* and self-reports present unique limitations. However, several current strategies increase the validity of self-reports in addition to physiological indicators, including bogus pipeline, construct validity models, peer ratings, parent and teacher ratings,

behavioral observations, and archival as well as secondary data sources. Finally, it is suggested that, based on current knowledge, the evidence for the validity of self-reports should not be underestimated.

Drug abuse prevention research findings and related technologies are not readily available to the prevention community and the general public. Certainly, more and better designed drug abuse prevention evaluation research has been initiated and completed in recent years. The interpretability of these research findings is also at a different level. However, much remains to be done in the area of transferring prevention technology to specific target groups. Schinke and Orlandi (this volume) suggest that technology transfer involves moving research findings from research into application. Kolbe and Iverson (1981) identify phases of implementing innovation and change, including mobilization for change that can be coupled with crisis, adoption of a commitment to a new program, implementation of the new practice, maintenance of the innovative program, and evolution of the program with changes. Schinke and Orlandi suggest four steps to increase the likelihood of technology transfer: (1) stick to the basics of conventional research designs; (2) replicate successful interventions, methodologies, and other results; (3) analyze the costs of interventions; and (4) strive for high-quality dissemination to enhance technology transfer.

CONSENSUS STATEMENTS RELATED TO PREVENTION RESEARCH METHODOLOGY

With the chapters in this publication as background, consensus recommendations were formulated by the meeting participants. The lively discussion during consensus development added to the unanimity of agreement. The following consensus statements are grouped into three areas, which were developed by the editors.

Modifying Existing Prevention Approaches

- Prevention evaluation studies should be theory based. This is essential for the further development of prevention as an area of science and for knowledge development. Unfortunately, past prevention evaluation efforts usually did not incorporate theoretical underpinnings but rather focused more on design issues and the evaluation aspects of exposure/control group comparisons. Prevention evaluation research should focus on empirically testing *a priori* theory.

- Prevention evaluation research has focused largely on the individual level. However, there also seems to be a developing interest in the social unit as

a level of analysis. Prevention evaluation research should incorporate, as appropriate, various social unit measures, including proximate ones such as the family and peer groups as well as larger social units such as schools, health care providers, and other institutional and environmental social units. Measuring the effects, both separate and combined, of these social units may add to the understanding of knowledge, attitude, and behavior shifts that are currently reflected within specific groups (i.e., senior high school students).

- Prevention evaluation research should be more innovative. For example, a better understanding of the informal system and its importance for influencing drug-seeking behavior as well as onset of drug use could enhance the impact of drug abuse prevention.

- The timing is right to develop prevention evaluation research studies that cut across disciplines and scientific areas. Using the best available data and information, social and behavioral prevention interventions should take into account multiple points of view and perspectives, including the basic sciences.

Integrating Epidemiological, Etiologic, and Intervention Research Methods

- Epidemiological and etiologic considerations should be incorporated into prevention evaluation research. In addition, prevention research should incorporate common methods that cut across etiology and epidemiology. The traditional separation may have hindered the development of new prevention interventions that subsequently might be tested in the real world. However, this separation does not add to the vitality and the interaction of findings to produce new and stronger prevention interventions that hold up under multiple conditions and replications. In other words, synergistic findings and cross-fertilization should be emphasized and stressed.

- A promising area for prevention evaluation research, from etiologic research, is behavioral genetics. For example, prevention research evaluations might incorporate biological markers to better understand behavioral genetic influences on the outcome of prevention interventions.

- Prevention evaluation research should incorporate efforts to better understand the maintenance and durability of prevention intervention effects over time. There needs to be a better understanding of the longitudinal effects of prevention interventions as well as efforts to maintain

these effects. In addition, evaluation research, as appropriate, should incorporate planned booster sessions to better understand the maintenance of prevention effects.

Expanding Prevention Research Utilization

- Additional emphasis should be placed on expanding minority research in the area of drug abuse prevention. The impact of drug abuse on minority communities is devastating, yet we know little about the types of prevention programs and initiatives that are effective in reducing the incidence and prevalence of drug abuse in these communities. Unfortunately, most current prevention interventions have been evaluated using majority populations and have not oversampled minority populations.

- Technical assistance should be available and related to drug abuse evaluation research and should be directed to local prevention programs so that they might evaluate their own prevention activities. In addition, technical assistance could help prevention programs to incorporate findings from process evaluation and possible outcome evaluation in refining their prevention interventions.

- Several multiphase studies should be developed that incorporate those basic prevention principles that have been proven to be effective from smaller scale prevention intervention evaluation studies (i.e., start early, incorporate multiple and time-phased interventions).

- There is a need to expand prevention evaluation research into nontraditional settings (i.e., neighborhoods, service organizations, policymaking bodies, and civic associations).

There was an extremely positive atmosphere among the meeting participants. The above recommendations are formulated to strengthen future research rather than to replace past and current prevention research efforts.

REFERENCES

Campbell, D.T., and Stanley, J.R. *Experimental and Quasi-Experimental Designs for Research.* Boston: Houghton Mifflin Company, 1963.

Hawkins, J.D.; Catalano, R.F.; Bridges, G.S.; Lake, L.; Gainey, R.; Murphy, T.; and Lishner, D. "A Risk-Based Analysis of Drug Abuse Prevention Strategies and Prospects." Unpublished manuscript, 1989.

Kolbe, L.J., and Iverson, D.C. Implementing comprehensive health education: Educational innovations and social change. *Health Educ Q* 8:57-80, 1981.

Lettieri, D.J.; Sayers, M.; and Pearson, H.W. *Theories on Drug Abuse: Selected Contemporary Perspectives.* Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1980.

Leukefeld, C.G. The role of the National Institute on Drug Abuse in drug abuse prevention. In: Donohew, L.; Sypher, H.; and Bukoski, W., eds. *Persuasive Communication and Drug Abuse Prevention.* Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Mark, M.M. Treatment implementation, statistical power, and internal validity. *Eval Rev* 7:543-549, 1983.

Rudner, R.S. *Philosophy of Social Sciences.* Englewood Cliffs, NJ: Prentice-Hall, 1966.

AUTHORS

Carl G. Leukefeld, D.S.W.
Director
Center for Drug and Alcohol Research
Department of Psychiatry
University of Kentucky College of Medicine
Annex 2, Room 206
Lexington, KY 40536-0080

William J. Bukoski, Ph.D.
Program Director
Intervention Research
Prevention Research Branch
Division of Epidemiology and Prevention Research
National Institute on Drug Abuse
Rockwall II, Suite 615
5600 Fishers Lane
Rockville, MD 20857

National Institute on Drug Abuse

Research monograph series

While limited supplies last, single copies of the monographs may be obtained free of charge from the National Clearinghouse for Alcohol and Drug Information (NCADI). Please contact NCADI also for information about availability of coming issues and other publications of the National Institute on Drug Abuse relevant to drug abuse research.

Additional copies may be purchased from the U.S. Government Printing Office (GPO) and/or the National Technical Information Service (NTIS) as indicated. NTIS prices are for paper copy; add $3 handling charge for each order. Microfiche copies are also available from NTIS. Prices from either source are subject to change.

Addresses are:

NCADI
National Clearinghouse for Alcohol and Drug Information
P.O. Box 2345
Rockville, MD 20852
(301) 468-2600
(800) 729-6686

GPO
Superintendent of Documents
U.S. Government Printing Office
Washington, DC 20402
(202) 275-2981

NTIS
National Technical Information Service
U.S. Department of Commerce
Springfield, VA 22161
(703) 487-4650

For information on availability of NIDA Research Monographs 1 through 24 (1975-1979) and others not listed, write to NIDA, Community and Professional Education Branch, Room 10A-54, 5600 Fishers Lane, Rockville, MD 20857.

25 BEHAVIORAL ANALYSIS AND TREATMENT OF SUBSTANCE ABUSE. Norman A. Krasnegor, Ph.D., ed.
GPO out of stock
NCADI out of stock
NTIS PB #80-112428/AS $31

26 THE BEHAVIORAL ASPECTS OF SMOKING. Norman A. Krasnegor, Ph.D., ed. (reprint from 1979 Surgeon General's Report on Smoking and Health)
GPO out of stock
NTIS PB #80-118755/AS $23

30 THEORIES ON DRUG ABUSE: SELECTED CONTEMPORARY PERSPECTIVES. Dan J. Lettieri, Ph.D.; Mollie Sayers; and Helen W. Pearson, eds.
GPO out of stock
NCADI out of stock
Not available from NTIS

31 MARIJUANA RESEARCH FINDINGS: 1980. Robert C. Petersen, Ph.D., ed.
GPO out of stock
NTIS PB #80-215171/AS $31

32 GC/MS ASSAYS FOR ABUSED DRUGS IN BODY FLUIDS. Rodger L. Foltz, Ph.D.; Allison F. Fentiman, Jr., Ph.D.; and Ruth B. Foltz, eds.
GPO out of stock
NCADI out of stock
NTIS PB #81-133746/AS $31

36 NEW APPROACHES TO TREATMENT OF CHRONIC PAIN: A REVIEW OF MULTIDISCIPLINARY PAIN CLINICS AND PAIN CENTERS. Lorenz K.Y. Ng, M.D., ed.
GPO out of stock
NCADI out of stock
NTIS PB #81-240913/AS $31

37 BEHAVIORAL PHARMACOLOGY OF HUMAN DRUG DEPENDENCE. Travis Thompson, Ph.D., and Chris E. Johanson, Ph.D., eds.
GPO out of stock
NCADI out of stock
NTIS PB #82-136961/AS $39

38 DRUG ABUSE AND THE AMERICAN ADOLESCENT. Dan J. Lettieri, Ph.D., and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report.
GPO out of stock
NCADI out of stock
NTIS PB #82-148198/AS $23

40 ADOLESCENT MARIJUANA ABUSERS AND THEIR FAMILIES. Herbert Hendin, M.D; Ann Pollinger, Ph.D.; Richard Ulman, Ph.D.; and Arthur Carr, Ph.D., eds.
GPO out of stock
NCADI out of stock
NTIS PB #82-133117/AS $23

42 THE ANALYSIS OF CANNABINOIDS IN BIOLOGICAL FLUIDS. Richard L. Hawks, Ph.D., ed.
GPO out of stock
NTIS PB #83-136044/AS $23

44 MARIJUANA EFFECTS ON THE ENDOCRINE AND REPRODUCTIVE SYSTEMS. Monique C. Braude, Ph.D., and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report.
GPO out of stock
NCADI out of stock
NTIS PB #85-150563/AS $23

45 CONTEMPORARY RESEARCH IN PAIN AND ANALGESIA, 1983. Roger M. Brown, Ph.D.; Theodore M. Pinkert, M.D., J.D.; and Jacqueline P. Ludford, M.S. eds. A RAUS Review Report.
GPO out of stock
NCADI out of stock
NTIS PB #84-184670/AS $17

46 BEHAVIORAL INTERVENTION TECHNIQUES IN DRUG ABUSE TREATMENT. John Grabowski, Ph.D.; Maxine L. Stitzer, Ph.D.; and Jack E. Henningfield, Ph.D., eds.
GPO out of stock
NCADI out of stock
NTIS PB #84-184688/AS $23

47 PREVENTING ADOLESCENT DRUG ABUSE: INTERVENTION STRATEGIES. Thomas J. Glynn, Ph.D.; Carl G. Leukefeld, D.S.W.; and Jacqueline P. Ludford, M.S., eds. A RAUS Review Report.
GPO out of stock
NCADI out of stock
NTIS PB #85-159663/AS $31

48 MEASUREMENT IN THE ANALYSIS AND TREATMENT OF SMOKING BEHAVIOR. John Grabowski, Ph.D., and Catherine Bell, M.S., eds.
GPO out of stock
NCADI out of stock
NTIS PB #84-145184/AS $23

50 COCAINE: PHARMACOLOGY, EFFECTS, AND TREATMENT OF ABUSE. John Grabowski, Ph.D., ed.
GPO Stock #017-024-01214-9 $4
NTIS PB #85-150381/AS $23

51 DRUG ABUSE TREATMENT EVALUATION: STRATEGIES, PROGRESS, AND PROSPECTS. Frank M. Tims, Ph.D., ed.
GPO out of stock NTIS PB #85-150365/AS $23

52 TESTING DRUGS FOR PHYSICAL DEPENDENCE POTENTIAL AND ABUSE LIABILITY. Joseph V. Brady, Ph.D., and Scott E. Lukas, Ph.D., eds.
GPO out of stock NTIS PB #85-150373/AS $23

53 PHARMACOLOGICAL ADJUNCTS IN SMOKING CESSATION. John Grabowski, Ph.D., and Sharon M. Hall, Ph.D., eds.
GPO out of stock NCADI out of stock
NTIS PB #89-123186/AS $23

54 MECHANISMS OF TOLERANCE AND DEPENDENCE. Charles Wm. Sharp, Ph.D., ed.
GPO out of stock NCADI out of stock
NTIS PB #89-103279/AS $39

55 PROBLEMS OF DRUG DEPENDENCE, 1984: PROCEEDINGS OF THE 46TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO out of stock NCADI out of stock
NTIS PB #89-123194/AS $45

56 ETIOLOGY OF DRUG ABUSE: IMPLICATIONS FOR PREVENTION. Coryl LaRue Jones, Ph.D., and Robert J. Battjes, D.S.W., eds.
GPO Stock #017-024-01250-5 $6.50 NTIS PB #89-123160/AS $31

57 SELF-REPORT METHODS OF ESTIMATING DRUG USE: MEETING CURRENT CHALLENGES TO VALIDITY. Beatrice A. Rouse, Ph.D.; Nicholas J. Kozel, M.S.; and Louise G. Richards, Ph.D., eds.
GPO out of stock NTIS PB #88-248083/AS $23

58 PROGRESS IN THE DEVELOPMENT OF COST-EFFECTIVE TREATMENT FOR DRUG ABUSERS. Rebecca S. Ashery, D.S.W., ed.
GPO out of stock NTIS PB #89-125017/AS $23

59 CURRENT RESEARCH ON THE CONSEQUENCES OF MATERNAL DRUG ABUSE. Theodore M. Pinkert, M.D., J.D., ed.
GPO out of stock NTIS PB #89-125025/AS $23

60 PRENATAL DRUG EXPOSURE: KINETICS AND DYNAMICS. C. Nora Chiang, Ph.D., and Charles C. Lee, Ph.D., eds.
GPO out of stock NTIS PB #89-124564/AS $23

61 COCAINE USE IN AMREICA: EPIDEMIOLOGIC AND CLINICAL PERSPECTIVES. Nicholas J. Kozel, M.S., and Edgar H. Adams, M.S., eds.
GPO out of stock NTIS PB #89-131866/AS $31

62 NEUROSCIENCE METHODS IN DRUG ABUSE RESEARCH. Roger M. Brown, Ph.D.; David P. Friedman, Ph.D.; and Yuth Nimit, Ph.D., eds.
GPO out of stock NCADI out of stock
NTIS PB #89-130660/AS $23

63 PREVENTION RESEARCH: DETERRING DRUG ABUSE AMONG CHILDREN AND ADOLESCENTS. Catherine S. Bell, M.S., and Robert Battjes, D.S.W., eds.
GPO out of stock NTIS PB #89-103287/AS $31

64 PHENCYCLIDINE: AN UPDATE. Doris H. Clouet, Ph.D., ed.
GPO out of stock NTIS PB #89-131858/AS $31

65 WOMEN AND DRUGS: A NEW ERA FOR RESEARCH. Barbara A. Ray, Ph.D., and Monique C. Braude, Ph.D., eds.
GPO Stock #017-024-01283-1 $3.50 NTIS PB #89-130637/AS $23

66 GENETIC AND BIOLOGICAL MARKERS IN DRUG ABUSE AND ALCOHOLISM. Monique C. Braude, Ph.D., and Helen M. Chao, Ph.D., eds.
GPO out of stock NCADI out of stock
NTIS PB #89-134423/AS $23

68 STRATEGIES FOR RESEARCH ON THE INTERACTIONS OF DRUGS OF ABUSE. Monique C. Braude, Ph.D., and Harold M. Ginzburg, M.D., J.D., M.P.H., eds.
GPO out of stock NCADI out of stock
NTIS PB #89-134936/AS $31

69 OPIOID PEPTIDES: MEDICINAL CHEMISTRY. Rao S. Rapaka, Ph.D.; Gene Barnett, Ph.D.; and Richard L. Hawks, Ph.D., eds.
GPO out of stock NTIS PB #89-158422/AS $39

70 OPIOID PEPTIDES: MOLECULAR PHARMACOLOGY, BIOSYNTHESIS, AND ANALYSIS. Rao S. Rapaka, Ph.D., and Richard L. Hawks, Ph.D., eds.
GPO out of stock NTIS PB #89-158430/AS $45

71 OPIATE RECEPTOR SUBTYPES AND BRAIN FUNCTION. Roger M. Brown, Ph.D.; Doris H. Clouet, Ph.D.; and David P. Friedman, Ph.D., eds.
GPO out of stock NTIS PB #89-151955/AS $31

72 RELAPSE AND RECOVERY IN DRUG ABUSE. Frank M. Tims, Ph.D., and Carl G. Leukefeld, D.S.W., eds.
GPO Stock #017-024-01302-1 $6 NTIS PB #89-151963/AS $31

73 URINE TESTING FOR DRUGS OF ABUSE. Richard L. Hawks, Ph.D., and C. Nora Chiang, Ph.D., eds.
GPO Stock #017-024-01313-7 $3.75 NTIS PB #89-151971/AS $23

74 NEUROBIOLOGY OF BEHAVIORAL CONTROL IN DRUG ABUSE. Stephen I. Szara, M.D., D.Sc., ed.
GPO Stock #017-024-01314-5 $3.75 NTIS PB #89-151989/AS $23

75 PROGRESS IN OPIOID RESEARCH. PROCEEDINGS OF THE 1986 INTERNATIONAL NARCOTICS RESEARCH CONFERENCE. John W. Holaday, Ph.D.; Ping-Yee Law, Ph.D.; and Albert Herz, M.D., eds.
GPO out of stock NCADI out of stock
Not available from NTIS

76 PROBLEMS OF DRUG DEPENDENCE, 1986: PROCEEDINGS OF THE 48TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO out of stock NCADI out of stock
NTIS PB #88-208111/AS $53

77 ADOLESCENT DRUG ABUSE: ANALYSES OF TREATMENT RESEARCH. Elizabeth R. Rahdert, Ph.D., and John Grabowski, Ph.D., eds.
GPO Stock #017-024-01348-0 $4 NCADI out of stock
NTIS PB #89-125488/AS $23

78 THE ROLE OF NEUROPLASTICITY IN THE RESPONSE TO DRUGS. David P. Friedman, Ph.D., and Doris H. Clouet, Ph.D., eds.
GPO out of stock NTIS PB #88-245683/AS $31

79 STRUCTURE-ACTIVITY RELATIONSHIPS OF THE CANNABINOIDS. Rao S. Rapaka, Ph.D., and Alexandros Makriyannis, Ph.D., eds.
GPO out of stock NTIS PB #89-109201/AS $31

80 NEEDLE SHARING AMONG INTRAVENOUS DRUG ABUSERS: NATIONAL AND INTERNATIONAL PERSPECTIVES. Robert J. Battjes, D.S.W., and Roy W. Pickens, Ph.D., eds.
GPO out of stock NTIS PB #88-236138/AS $31

81 PROBLEMS OF DRUG DEPENDENCE, 1987: PROCEEDINGS OF THE 49TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO Stock #017-024-01354-4 $17 NTIS PB #89-109227/AS Contact NTIS for price

82 OPIOIDS IN THE HIPPOCAMPUS. Jacqueline F. McGinty, Ph.D., and David P. Friedman, Ph.D., eds.
GPO out of stock NTIS PB #88-245691/AS $23

83 HEALTH HAZARDS OF NITRITE INHALANTS. Harry W. Haverkos, M.D., and John A. Dougherty, Ph.D., eds.
GPO out of stock NTIS PB #89-125496/AS $23

84 LEARNING FACTORS IN SUBSTANCE ABUSE. Barbara A. Ray, Ph.D., ed.
GPO Stock #017-024-01353-6 $6 NTIS PB #89-125504/AS $31

85 EPIDEMIOLOGY OF INHALANT ABUSE: AN UPDATE. Raquel A. Crider, Ph.D., and Beatrice A. Rouse, Ph.D., eds.
GPO Stock #017-024-01360-9 $5.50 NTIS PB #89-123178/AS $31

86 COMPULSORY TREATMENT OF DRUG ABUSE: RESEARCH AND CLINICAL PRACTICE. Carl G. Leukefeld, D.S.W., and Frank M. Tims, Ph.D., eds.
GPO Stock #017-024-01352-8 $7.50 NTIS PB #89-151997/AS $31

87 OPIOID PEPTIDES: AN UPDATE. Rao S. Rapaka, Ph.D., and Bhola N. Dhawan, M.D., eds.
GPO Stock #017-024-01366-8 $7 NTIS PB #89-158430/AS $45

88 MECHANISMS OF COCAINE ABUSE AND TOXICITY. Doris H. Clouet, Ph.D.; Khursheed Asghar, Ph.D.; and Roger M. Brown, Ph.D., eds.
GPO Stock #017-024-01359-5 $11 NTIS PB #89-125512/AS $39

89 BIOLOGICAL VULNERABILTY TO DRUG ABUSE. Roy W. Pickens, Ph.D., and Dace S. Svikis, B.A., eds.
GPO Stock #017-022-01054-2 $5 NTIS PB #89-125520/AS $23

90 PROBLEMS OF DRUG DEPENDENCE, 1988: PROCEEDINGS OF THE 50TH ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO Stock #017-024-01362-5 $17

91 DRUGS IN THE WORKPLACE: RESEARCH AND EVALUATION DATA. Steven W. Gust, Ph.D., and J. Michael Walsh, Ph.D., eds.
GPO Stock #017-024-01384-6 $10 NTIS PB #90-147257/AS $39

92 TESTING FOR ABUSE LIABILITY OF DRUGS IN HUMANS. Marian W. Fischman, Ph.D., and Nancy K. Mello, Ph.D., eds.
GPO Stock #017-024-01379-0 $12 NTIS PB #90-148933/AS $45

93 AIDS AND INTRAVENOUS DRUG USE: FUTURE DIRECTIONS FOR COMMUNITY-BASED PREVENTION RESEARCH. C.G. Leukefeld, D.S.W.; R.J. Battjes, D.S.W.; and Z. Amsel, D.Sc., eds.
GPO Stock #017-024-01388-9 $10 NTIS PB #90-148941/AS $39

94 PHARMACOLOGY AND TOXICOLOGY OF AMPHETAMINE AND RELATED DESIGNER DRUGS. Khursheed Asghar, Ph.D., and Errol De Souza, Ph.D., eds.
GPO Stock #017-024-01386-2 $11 NTIS PB #90-148958/AS $39

95 PROBLEMS OF DRUG DEPENDENCE, 1989: PROCEEDINGS OF THE 51ST ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.
GPO Stock #017-024-01399-4 $21 NTIS PB #90-237660/AS $67

96 DRUGS OF ABUSE: CHEMISTRY, PHARMACOLOGY, IMMUNOLOGY, AND AIDS. Phuong Thi Kim Pham, Ph.D., and Kenner Rice, Ph.D., eds.
GPO Stock #017-024-01403-6 $8 NTIS PB #90-237678/AS $31

97 NEUROBIOLOGY OF DRUG ABUSE: LEARNING AND MEMORY. Lynda Erinoff, Ph.D., ed.
GPO Stock #017-024-01404-4 $8 NTIS PB #90-237686/AS $31

98 THE COLLECTION AND INTERPRETATION OF DATA FROM HIDDEN POPULATIONS. Elizabeth Y. Lambert, M.S., ed.
GPO Stock #017-024-01407-9 $4.75 NTIS PB #90-237694/AS $23

99 RESEARCH FINDINGS ON SMOKING OF ABUSED SUBSTANCES. C. Nora Chiang, Ph.D., and Richard L. Hawks, Ph.D. eds.
GPO Stock #017-024-01412-5 $5

100 DRUGS IN THE WORKPLACE: RESEARCH AND EVALUATION DATA. VOL. II. Steven W. Gust, Ph.D., and J. Michael Walsh, Ph.D., eds.

101 RESIDUAL EFFECTS OF ABUSED DRUGS ON BEHAVIOR. John W. Spencer, Ph.D., and John J. Boren, Ph.D., eds.
GPO Stock #017-024-01426-5 $6

102 ANABOLIC STEROID ABUSE. Geraline C. Lin, Ph.D., and Lynda Erinoff, Ph.D., eds.
GPO Stock #017-024-01425-7 $8

103 DRUGS AND VIOLENCE: CAUSES, CORRELATES, AND CONSEQUENCES. Mario De La Rosa, Ph.D.; Elizabeth Y. Lambert, M.S.; and Bernard Gropper, Ph.D., eds.
GPO Stock #017-024-01427-3 $9

104 PSYCHOTHERAPY AND COUNSELING IN THE TREATMENT OF DRUG ABUSE. Lisa Simon Onken, Ph.D., and Jack D. Blaine, M.D., eds.
GPO Stock #017-024-01429-0 $4

105 PROBLEMS OF DRUG DEPENDENCE, 1990: PROCEEDINGS OF THE 52ND ANNUAL SCIENTIFIC MEETING, THE COMMITTEE ON PROBLEMS OF DRUG DEPENDENCE, INC. Louis S. Harris, Ph.D., ed.

106 IMPROVING DRUG ABUSE TREATMENT. Roy W. Pickens, Ph.D.; Carl G. Leukefeld, D.S.W.; and Charles R. Schuster, Ph.D., eds.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1991-294-803