

*EXTENDING THE CONCEPT OF SOCIAL VALIDITY:
BEHAVIOR ANALYSIS FOR DISEASE
PREVENTION AND HEALTH PROMOTION*

RICHARD A. WINETT, JOHN F. MOORE, AND EILEEN S. ANDERSON

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

A broader definition of social validity is proposed wherein a socially valid behavior-change intervention is directed to a problem of verifiable importance, the intervention is valued and used appropriately by designated target groups, and the intervention as used has sufficient behavioral impact to substantially reduce the probability of the problem's occurrence in target populations. The verifiable importance of a problem is based on epidemiological data, and the value and appropriate use of an intervention are enhanced through the use of conceptual frameworks for social marketing and behavior change and considerable formative and pilot research. Behavioral impact is assessed through efficacy and effectiveness studies. Thus, the social validity of a behavior-change intervention is established through a number of interactive, a priori steps. This approach to defining social validity is related to critical analysis and intervention issues including individual and population perspectives and "top-down" and "bottom-up" approaches to intervention design. This broader definition of social validity is illustrated by a project to reduce the risk of HIV infection among adolescents. Although the various steps involved in creating socially valid interventions can be complicated, time-consuming, and expensive, following all the steps can result in interventions capable of improving a nation's health.

DESCRIPTORS: social validity, epidemiology, social marketing, AIDS prevention, health promotion

According to Wolf (1978), the users (or clients) of an intervention program should evaluate the significance of the goals of that intervention, the social appropriateness of its procedures, and the social importance of its effects. The results of such evaluations are measures of "social validity." As stated by Wolf: "... if we aspire to social importance, then we must develop systems that allow our consumers to provide us feedback about how our applications relate to their values, to their reinforcers" (p. 475).

Presumably, interventions that show both strong behavioral outcomes and positive social validity

ratings are the ones more likely to be purposely disseminated and adopted. Interventions with clear evidence of behavioral outcome, but with low social validity, may be purposely disseminated but will have a lower probability of being adopted (Stolz, 1981). The last point indicates that social validity is directly related to such broader areas as the dissemination of innovations (Rogers, 1983) and appropriate community technologies (Fawcett, Seekins, Whang, & Muiu, 1984).

The early exposition of social validity (Wolf, 1978) was apparently presented as a resolution of the conflict between objective behavioral measures of outcome and more subjective post hoc measures of value, importance, and utility by consumers. It was decided that post hoc subjective consumer measures can be an admissible part of behavior-analytic work because these measures are necessary to refine and validate interventions. This discussion offers a broader perspective of social validity. From our perspective, social validity should be a process using epidemiological data and a series of interactive a priori steps, including frequent contact with potential users of an intervention and the application of both objective and subjective measures. Thus, we

Work on this paper was supported by grants from the National Institute of Mental Health and the National Cancer Institute. The project reported was a cooperative project between the Center for Research in Health Behavior and the Learning Resources Center. We acknowledge the helpful feedback on an earlier version of this paper from David Altman and Scott Geller. We also acknowledge the contributions of Lee Hite, Tamara Neubauer, Richard Hook, Jana Wagner, David Lombard, Kathleen Sikkema, Deborah Webster, David Taylor, Michael Leahy, and Bruce Walker to project-related work discussed in this paper. Correspondence and reprint requests should be sent to Richard A. Winett, Center for Research in Health Behavior, Department of Psychology, Virginia Tech, Blacksburg, Virginia 24061.

have prepared this discussion article to (a) broaden the boundaries of social validity, (b) articulate measures and processes involved in assessing social validity, and (c) provide examples of the interactive steps involved in designing, refining, and evaluating a socially valid intervention program.

DEFINITION AND BOUNDARIES

A socially valid behavior-change intervention is directed to a problem of verifiable importance, in which intervention is valued and used appropriately by designated target groups and has sufficient behavioral impact to reduce substantially the probability of the problem's occurrence among the target group(s). This definition is broad and attempts to include evaluations of "efficacy," "use," and "impact." Terms such as "verifiable importance," "used appropriately and valued," "sufficient behavioral impact," and "reduce the probability of occurrence" need to be, if not strictly defined, at least discussed and illustrated. The discussion and illustrations in this and later sections focus primarily on disease prevention through the promotion of health behaviors.

Verifiable Importance: Epidemiology

Verifiable importance relates to epidemiology, a field concerned with ". . . the origin and spread of disease in human populations" (Palinkas, 1985, p. 49). Epidemiology is the focus of the public health domain on studies assessing relationships among physiological, social, psychological, and environmental events and the incidence and prevalence of chronic and infectious diseases. The *incidence rate* of a disease is the number of new cases of a disease within a specified population over a specified time period. The *prevalence rate* measures the total number of cases of a disease at one moment or during a specified time interval (see Palinkas, 1985, for more complete definitions).

As described by Bloom (1988), there are two general approaches taken to investigate the relationship between a suspected causal agent and a disorder. In retrospective studies (case-control studies), a comparison is made between a group of individuals with the disorder (e.g., coronary heart

disease) and a matched group of individuals without the disorder. An analysis of the prior history of the two groups could determine a factor (e.g., sedentary lifestyle as defined by mean caloric expenditure in leisure activities) that occurs more frequently in the group with the disorder. Even large retrospective studies are not expensive to conduct, and many factors can be analyzed in one study. Therefore, retrospective studies are usually done first in an epidemiological research program.

Prospective studies (cohort studies) identify matched groups of persons with (e.g., sedentary lifestyle) or without (e.g., active lifestyle) a presumed risk factor for a disorder. The two groups are followed for an appropriate period of time (e.g., age 40 to 60) to determine the rates of the disorder in each group. Therefore, prospective studies can determine the actual rate of risk associated with a factor (e.g., sedentary lifestyle) and a number of diseases (e.g., coronary heart disease, colon cancer) or disorders (e.g., mental health problems). However, these studies are expensive and take years to conduct. Thus, retrospective studies are generally conducted first to establish relationships between diseases and disorders and potential risk factors. Prospective studies can provide convincing evidence to verify relationships found in retrospective studies and to calculate actual risk functions.

Bloom (1988) described six criteria used by epidemiologists to establish a causal relationship between a presumed risk factor and a disease or disorder. The *consistency of association* refers to similar findings emerging in studies conducted by different investigators at different times and using different methods, sites, and populations. The greater the consistency of association across studies, the greater the confidence in the causal relationship between a risk factor and a disease or disorder.

The *strength of association* involves comparisons of morbidity and mortality rates for persons with or without the presumed risk factor. The rates are used in a ratio that becomes the relative risk ratio. For example, at a 20-year follow-up there may be 20 cases of heart disease in a group of 100 inactive persons, but only four cases of heart disease in a matched group of 100 active persons. The

relative risk ratio ($20\%/4\% = 5$) shows the strength of association, which in this example is considered high (i.e., very strong association; Jeffrey, 1989).

The *specificity of association* refers to whether or not the presumed risk factor is associated with one disease or disorder. Specificity increases arguments for causality, but lack of specificity does not rule out a factor as a causal agent. For example, smoking is related to many diseases and is specifically related to lung cancer (Kuller, Meilahn, Townsend, & Weinberg, 1982).

The *temporal relationship of association* refers to whether or not exposure to a risk factor precedes the disease or disorder. Prospective studies can verify temporal relationships and add to the credibility of a presumed risk factor.

The *coherence of association* involves the degree of agreement about the risk factor and the known facts about the disease. Two key criteria are whether or not exposure to the risk factor and the development of the disease or disorder make biological sense, and whether or not greater exposure to the presumed risk factor increases the probability of the occurrence of the disease or disorder (called the dose-response relationship). Smoking and the occurrence of lung cancer is an association that makes biological sense and exhibits a dose-response relationship (Kuller et al., 1982).

Preventive clinical trials can provide final, relatively definitive evidence of a causal relationship between a presumed risk factor and a disease or disorder. In controlled clinical trials, study participants are randomly assigned to conditions. Typically, the experimental group receives an intervention that greatly reduces or eliminates the risk factor, whereas the control group receives no intervention or a very minimal intervention. The subsequent reduction in the rates of a specific disorder in the experimental group but not the control group supports the causal relationship between the risk factor and the disorder.

It is often difficult, however, to find sets of epidemiological data for specific diseases that fit all six criteria. One clear exception is cigarette smoking (Kuller et al., 1982), with others in the area of injury control (e.g., Robertson, 1986). Although

there is some disagreement in the field, epidemiological evidence linking dietary factors (total fat, saturated fat), high blood cholesterol, and coronary heart disease appear to fulfill closely the six criteria for establishing a relationship between a risk factor and a disease (Kwiterovich, 1989; Trevisan et al., 1990). Other causative factors and potential risk behaviors of interest (dietary factors and cancer [Committee on Diet and Health, 1989; Hursting, Thornquist, & Henderson, 1990; Koop, 1988]; number of sexual partners, type of drug use, and HIV infection [Turner, Miller, & Moses, 1989]) often have compelling but less complete data or, at times, inconsistent data. For example, we are unaware of studies showing a gradient (dose-response) relationship between amount of fiber in the diet and colon cancer. Instead, the most convincing studies are cross-cultural retrospective studies (e.g., Hursting et al., 1990). One brief prevention trial did not verify a relationship between a large reduction of fat in the diet for experimental participants and reduced rates of breast cancer (Henderson et al., 1990), although retrospective study data suggested a relationship between dietary fat and breast cancer (Koop, 1988).

Relative risk is an important statistic to be considered in assessing verifiable importance. As discussed above, relative risk is a specific measure of the strength of association between a risk factor and morbidity and mortality rates. It is the proportion of two rates (e.g., the rate of heart disease in a group of smokers compared to the rate of heart disease in a matched group of non-smokers as derived from a prospective study).

A good example of recent research on relative risk is in the area of physical fitness. The relationship between cardiovascular disease and leisure activity (measured as estimated weekly caloric expenditure in leisure activities) has been documented previously (Paffenbarger, Hyde, Wing, & Hsieh, 1986). However, the relationship between specific fitness levels and the risk of premature death was unknown until recently.

Blair et al. (1989) followed 10,224 men and 3,120 women for about 8 years after they had been given a maximal treadmill exercise test as part of

a thorough preventive medical examination. Based on standards for the treadmill test, the men and women were placed in one of five fitness categories (lowest to highest). All-cause mortality rates were calculated separately for men and women from each fitness category. The mortality rates for the most fit category for men and women separately were used as the denominator in risk ratios for men and women. The numerator was the all-cause mortality rate for each separate men's and women's fitness category. For men, the relative risk of the lowest fitness category was 3.44. However, the relative risk of the next-to-lowest fitness category was 1.37. For women the relative risk for the lowest category was 4.65, for the next to lowest, 2.42, and for the third lowest, 1.46. The relative risk ratios show the strength of the association between fitness level and all-cause mortality. If degrees of fitness are considered as degrees of exposure to a risk factor (low fitness equals high exposure), there is also evidence for a dose-response relationship. However, the data for both men and women indicated that relative risk was substantially reduced at *moderate* fitness levels. These moderate fitness levels equate to only 30 to 60 min per day of brisk walking (or another modest aerobic activity) and are readily achievable by most children and adults (see Cooper, 1988).

In this case, relative risk data support the verifiable importance of physical fitness as a focus for reducing mortality rates, and the Blair *et al.* (1989) study also suggested guidelines and goals for intervention. If formerly sedentary, unfit individuals can be trained and persuaded to maintain a modest (though definable) amount of aerobic activity, their all-cause mortality risk can be greatly reduced. Similar data are available for total cholesterol level, low-density lipoprotein cholesterol level, and coronary heart disease risk (Kwiterovich, 1989), resulting in specific dietary guidelines (reduced total fat and saturated fat) to reduce heart disease risk.

The examples using fitness level, all-cause mortality, lipid profile, and coronary heart disease have, however, oversimplified the process of using epidemiological data to define verifiable importance. This is the case even when epidemiological studies in an area of concern are well developed. This is

because relative risk ratios pertain to individual risk. Relative risk ratios do not indicate the degree of risk for large populations nor do they suggest how resources may be best used to reduce risk on a population basis.

Population-attributable risk. Jeffrey (1989) discussed the inappropriateness of using relative risk to prioritize health problems and enact interventions. The more relevant concept is population-attributable risk, "... the number of excess cases of disease in a population that can be attributed to a particular risk factor" (Jeffrey, 1989, p. 1194). Population-attributable risk is based on the absolute risk of getting a certain disease, the population prevalence of the risk behavior, and the relative risk ratio. This statistic provides an estimate of the societal burden of unnecessary morbidity and mortality and the number of potential lives saved if an entire population stopped a particular risk behavior.

Importantly, a high relative risk ratio may not be that important from a societal or population level of analysis. This is the case when the number of persons at risk is small and/or the risk factor is of limited prevalence. For example, rock climbing may result in very high relative risk, but few people can climb rocks, and far fewer rock climbers will actually suffer serious injury or death from climbing.

Risk behaviors of low or moderate relative risk are often of particular interest from a societal perspective when engagement in the risk behavior is widespread (i.e., high prevalence) and when the risk behavior is associated with a variety of diseases. For example, measures of physical activity, in contrast to actual physical fitness (see Blair *et al.*, 1989), have been moderately associated with coronary heart disease, hypertension, non-insulin-dependent diabetes mellitus, osteoporosis, obesity, colon cancer, and mental health problems (Harris, Caspersen, DeFries, & Estes, 1989). In addition, it is estimated that over 80% of the population is sufficiently inactive to be at risk for these various diseases and disorders (Harris *et al.*, 1989). Thus, using the concept of population-attributable risk makes activity promotion a prime candidate for large-scale interventions to reduce morbidity and mortality.

The overall aim of choosing target problems on the basis of available epidemiological data is to rationalize the priority of problems. The concept of population-attributable risk is central to this prioritization. However, it is easy to see, even from the brief examples in this section, that it may often be difficult to rally our society to support the alteration of pervasive, seemingly innocuous, risk behaviors such as a sedentary lifestyle. Moreover, it would be naive to believe that problems, guidelines, and goals selected by "experts" will be the ones always favored by target groups or communities, or that objective data will be the only basis for deciding priorities and suggesting guidelines and goals for interventions.

However, as health-behavior and public health experts, we should try to bring a degree of rationality to problem focus and target behaviors. At a minimum, epidemiological data provide one basis for defining verifiable importance and for prioritizing problems. Note, however, that although epidemiological data can suggest guidelines and goals for interventions, epidemiological data do not suggest how to design an intervention program for large-scale acceptability, applicability, and impact.

Intervention Design: Social Marketing Framework and Related Issues

Epidemiological data serve as one basis for rationalizing the need for certain behavior-change interventions. These data can also suggest the guidelines and goals for target behavior change. For example, epidemiological data can point to the amount of aerobic exercise (calories expended per week in activity) or amount of fat in the diet (30% or less of calories from fat) appearing to reduce risks of heart disease. However, these data say little about how the intervention is presented to target groups, the level of intervention (e.g., individual, group, entire communities), or the format of the intervention. For example, key tasks can involve making exercise and diet plans both appealing to target groups and supportable by environmental contingencies so that the plans are appropriately followed.

Issues and questions regarding intervention de-

sign have been most frequently addressed by the field of social marketing (cf. Geller, 1989; Kotler & Roberts, 1989; Manoff, 1985). Social marketing adapts the framework and methods of commercial marketing to the realm of promoting ideas, social causes, and health-behavior changes. In the present context, we subsume under the term *social marketing* a framework and specific variables useful for conceptualizing and designing interventions and formative research, pilot research, and prototype testing strategies useful for further designing and refining an intervention (Winett, 1986). The framework includes four key variables: the *product* (i.e., the program, intervention, or system), and its *price* (including social and psychological costs), *place* (where the product is available and how it is distributed), and *promotion* (how the product is disseminated). *Positioning*, which means the particular niche of the product as distinct from similar products, is another term often used.

A change in one marketing variable usually results in changes in other marketing variables. For example, to reduce the time and effort involved in exercise programs at recreational centers, a new exercise program may be put on a videotape program for the home. The product and its place have been repositioned to reduce the price (i.e., response cost). Notice, however, that other costs increase (one-time payment of the video program, effort for rental), another promotional strategy may be needed, and the videotape has to be differentiated from the many other available exercise tapes. Further (and as discussed later), it is not clear if a program to be used alone in the home will be as effective as a program originally designed to be used with a group of exercisers and a trained exercise leader.

In addition to marketing variables and considerations, intervention design requires conceptual input. Conceptual input includes epidemiological data suggesting target groups, guidelines and goals for behavior change (Geller, 1989) and supportive environmental changes (Jeffrey, 1989; Winett, King, & Altman, 1989). The danger in minimizing conceptual input is the potential for developing interventions with popular appeal and ready distribution channels that nevertheless may be ineffective. The

particular conceptual processes and frameworks for creating disease-prevention and health-promotion interventions have been detailed in Winett *et al.* (1989) and Geller (1989) and will not be elaborated here.

In addition to promotional and conceptual consideration, the design of an intervention valued and used appropriately by target groups involves an a priori series of interactive formative, pilot, and prototype research steps. These steps are taken *before* the product is in its final form not only to determine acceptance rates but also to modify the proposed intervention so it can be optimally used by members of a target group. These steps are described in the next section. First, the general orientation of social marketing and related issues are discussed.

Social marketing can be construed as a "top-down" behavioral engineering approach, in contrast to "bottom-up" citizen collaboration models (Fawcett, 1990). With social marketing, the goals and minimal guidelines for an intervention have been decided by "experts." Citizen collaboration approaches enable citizens to choose their goals and intervention programs. However, rather than being necessarily contrasting or competing approaches, both approaches may be considered in the same program. Through a series of formative and pilot research steps, social marketing adherents are likely to modify the goals and emphases of an intervention (e.g., content of a video program) according to feedback from specific target groups and the overall community. It is possible that an intervention program developed by experts is not in a suitable form or even perceived as important by target groups in the community, but critical feedback should lead social marketers to reprioritize, reformulate, and restructure their plans.

Adherents of citizen collaboration models, however, are not entirely directed by members of target groups and the overall community. Surely, citizen collaboration adherents have some favored goals (e.g., teaching citizens how to use their collective power to influence local governmental policies) and intervention sites (e.g., grass-roots organization) of their own. Likewise, it is doubtful that these adherents are willing to lend their expertise and service

to any goal or any intervention program. Further, neither the involvement of citizens, experts, nor high versus low technology appears to be the exclusive province of either approach.

As noted earlier, both top-down and bottom-up approaches should be informed by conceptual and strategic inputs (including epidemiological data). The use of these inputs may be very apparent for the expert-driven, top-down approach, but may not be so obvious for citizens' work on problems and interventions. Although much can be gained from the development of interpersonal and political skills within grass-roots community organizations, even some of the staunchest adherents to the bottom-up approach agree that choosing goals and strategies may require expert guidance (e.g., see the biography of Saul Alinsky in Horwitt, 1989).

Thus, both top-down and bottom-up approaches involve experts who can influence goals and intervention design. Perhaps citizens are involved more passively in top-down approaches and more actively and collaboratively in bottom-up approaches (Wandersman, 1981). However, rather than being mutually exclusive, these approaches may be mutually beneficial, each informing the other. Moreover, the formative and pilot research steps of both approaches can overlap.

Intervention Design: Formative and Pilot Research to Increase Value and Appropriate Use

Formative research steps can take many forms. Rather than following a rigorous scientific method, the formative research steps are more intuitive but contain some logic in the sequence of steps (Winett, Altman, & King, 1990). However, the key questions of formative research involve the marketing variables described earlier, and formative research uses established information-gathering procedures. The steps are interactive because information obtained in one step suggests information needed in a subsequent step. Further, disconfirming information will often suggest returning to an earlier step.

Using exercise promotion as an example, for-

mative research steps may involve the identification of:

1. Target groups who exercise minimally but are interested in starting to exercise (Prochaska & DiClemente, 1983). This step may involve telephone or mail surveys.

2. Types of exercise protocols acceptable to particular target groups, and perceived personal, social, and environmental barriers to exercise (Hovell et al., 1989). This step may use more detailed, face-to-face structured interviews.

3. Several optimal exercise protocols that meet physiological guidelines for later presentation to members of target groups. The protocols would presumably also incorporate a number of key behavior-change strategies. This step usually involves "focus groups," small semistructured groups that provide opinions and verbal and written feedback on potential products.

Most formative research and the entire social marketing process appear to end at this point (Winnett, 1986). The option of returning to an earlier step (e.g., conduct more interviews to develop another protocol for testing in focus groups) is available and sometimes taken, but typically the next step is an intervention trial. This is unfortunate because the information-gathering steps just described rely exclusively on verbal reports. Rarely are facsimile interventions and observable behavioral measures used as the next testing step (i.e., a "behavioral social marketing" approach; Geller, 1989). Moreover, there are intermediate steps between formative research and putting a product into the market place that are done with commercial products. For example, the next steps in marketing a commercial product can include observations of customers using samples of the product or time-limited market tests in one or several locations.

For social products, the intermediate steps of facsimile design and pilot testing need to be included as part of an overall social marketing approach. These steps provide behavioral data for the redesign and refinement of products so the products work effectively. A facsimile is a simple, working prototype of the product. It is promoted to target groups in much the same way as the eventual

product. However, the facsimile usually provides considerable opportunity for unobtrusive observations of users with the product and direct interaction with the target group members while they use the product. Clearly, these observations and interactions present opportunities to refine the product to fit the target group better.

After the facsimile is refined, other facsimile tests may be conducted. At some point, feedback data are sufficiently positive to conduct a series of pilot studies. The sine qua non of pilot tests is to deliver and evaluate the intervention in the same way as a subsequent larger field test and to maintain the flexibility to return to another facsimile design (or even earlier stages of formative research) if the pilot study's outcomes are negative or only marginally positive. If the pilot study cannot be operationalized as the full-scale test and/or the option does not exist to modify the intervention, the pilot study serves, at best, as a dress rehearsal. Even if the procedures work smoothly, it is difficult to justify putting into full-scale operation an intervention that has been poorly evaluated, or worse, does not produce intended effects.

Although the formative and pilot research steps may appear quite logical, we note that until recently many large-scale social and behavior-change programs followed few of these steps (e.g., Winnett et al., 1990). Following these steps a priori does not assure an effective intervention program. However, the entire process should increase the probability that an intervention will be valued, used appropriately, and effective. In other words, the process increases the intervention's social validity.

Sufficient Behavioral Impact

The criterion of sufficient behavioral impact has several dimensions. These dimensions include *efficacy* and *effectiveness* at producing desired outcomes. Additionally, the scope and intensity of an intervention should match the degree of effectiveness and the ability of the intervention to reduce risk appreciably.

Guidelines for conducting risk-reduction research (e.g., Greenwald, Sondik, & Lynch, 1986) distinguish between efficacy and effectiveness. Ef-

ficacy refers to significant (statistical and clinical) behavioral outcomes within the context of initial testing of an intervention under relatively optimal conditions. Clearly, efficacy testing follows the pilot testing stage discussed earlier. On the other hand, effectiveness refers to significant behavioral impact under more typical, real-world conditions.

Efficacy testing. Initial testing usually involves the participation of homogeneous groups of people (e.g., as defined by socioeconomic status, age, and health status) studied within relatively circumscribed situations (e.g., several private physicians' practices) and with rigorous experimental methods (e.g., random assignment to "usual care" and "enhanced care" conditions). The major questions for efficacy tests are: Does this intervention work under optimal conditions, and if the intervention works, how well (or how much) does it work?

Efficacy testing is a critical first step. If an intervention does not produce intended results under optimal (contrived) conditions, there is little reason to expect the intervention to work under less optimal (natural) conditions. In fact, most of the articles appearing in research journals can be considered efficacy tests.

For an intervention to be socially valid, however, it must produce the desired effects under more typical (i.e., real-world) circumstances. For example, parents and teens may learn how to use problem-solving strategies at home when the family receives six 2-hour sessions delivered by a highly trained therapist in the family's home. However, the frequency and intensity of this intervention for large-scale application in nonclinical populations appear impractical. A more socially valid intervention might require a family-skills manual or videotape program.

This point suggests that the process of developing an effective intervention program is also guided by delivery system and cost factors. Questions pertaining to different delivery systems and their concomitant costs, goals, and expected outcomes are complex ones discussed at length elsewhere (Winnett *et al.*, 1989) and are only summarized below.

Effectiveness testing. The effectiveness test is an evaluation of an intervention under conditions du-

plicating those in which the intervention will be used. Such tests in the health-promotion and disease-prevention domain may involve many different kinds of patients in large health maintenance organizations (HMOs). Presumably, conditions are similar across HMOs. Doctors and nurses are busy, and many patients are also pressed for time. For example, nutrition interventions for hypertension, hypercholesterolemia, and overweight must be effective in a situation in which diverse patients have contact with a nurse or physician for as little as 15 minutes every 2 weeks (Fisher, 1989). An evaluation study showing a nutrition intervention to be effective under these conditions would meet one important criterion for social validity.

Most behavioral health interventions also require maintenance of behavior change. For nutrition and exercise interventions to be effective, for example, they need to be followed for long periods of time—indeed, for a lifetime. This is clearly the case when nutrition and exercise interventions are used to reduce hypertension, hypercholesterolemia, and overweight. Thus, where an efficacy test may be short term, an effectiveness test is usually longer term, with follow-up phases often covering many years.

Besides maintaining behavior change, a socially valid intervention should make a difference in risk categorization. For example, a nutrition intervention should enable individuals with "borderline high" blood pressure to reach the "normal" category or reduce total cholesterol from the "borderline high" range to a "normal" range. Unfortunately, it is rare that a single intervention program can accomplish this goal when delivered under real-world conditions. A nutrition intervention may theoretically have the potential to help most individuals reach blood pressure or cholesterol goals. However, as discussed earlier, real-world conditions at HMOs often limit medical staff and patient contact and, thus, limit effectiveness.

If an intervention cannot, as typically delivered, alter an individual's risk categorization, it should at least reduce risk in proportion to such factors as the ratio of intervention agents to program participants, duration of agent-to-participant contact, the duration of the behavior-change strategy, the degree

of participant involvement in the program, and the amount of social support developed to reinforce behavior change (Geller, 1991). These factors typically differentiate primary and secondary prevention efforts, with these efforts having different goals and operations. Many behavioral interventions represent secondary prevention efforts. Secondary prevention entails the treatment of persons already manifesting early symptoms of a disorder or persons at high risk for a disorder. Thus, individuals at risk are usually identified through screening; intensive and sustained treatment is applied to alter the course of a disorder, reduce its severity or likelihood of reoccurrence or, where possible, decrease risk of the disorder itself. Thus, secondary prevention is often a case-finding, clinical approach. One example of secondary prevention is the aggressive treatment of borderline hypertension with diet and exercise (Keleman, Effron, Valenti, & Stewart, 1990).

Primary prevention is usually population-based. The goal of primary prevention is to reduce the incidence of a disease or disorder in the population. A large-scale media and supermarket campaign that exhorts people to decrease fat and increase fiber in their diets is an example of a primary prevention intervention to reduce risks of cancer. If the campaign was successful, in 10 to 15 years the annual incidence of dietary-related cancers in the population would be reduced significantly (Greenwald et al., 1986).

Thus, most often, the outcome expectations for primary and secondary interventions are quite different. With primary prevention at the individual level of analysis, the expectation (compared to a clinical approach) is modest behavior change. Expectations are modest because factors usually associated with a greater degree of behavior change (e.g., sustained intervention-agent and participant contact; Geller, 1991) cannot usually be maximized in large-scale primary prevention efforts. However, more modest behavior changes can result in very meaningful reductions of risk and decreases in morbidity and mortality rates at the population level because so many persons can be affected (Farquhar et al., 1990). With secondary prevention, outcome expectations are for considerable behavior

change because the interventions are often intensive clinical treatments. For example, one-to-one dietary treatment can reduce total cholesterol levels by 10% to 15% and thereby decrease coronary heart disease risk in treated persons by 20% to 30% (Kwiterovich, 1989). At the individual level, these outcomes are impressive. However, the caveat is that the intensity and cost of these interventions are substantial and they typically cannot be applied on a large scale, thereby limiting population effects.

Evaluating the relative cost effectiveness and cost-benefit relationship of primary and secondary prevention strategies is a highly complex task and is beyond the scope of this discussion. However, the points in this section show that an effective intervention is defined by its behavioral outcomes with regard to real-world circumstances, type and scope of the intervention, and absolute risk reduction at the individual and population levels.

Including criteria for effectiveness alters the concept of social validity in another way. Using these criteria, it does not seem appropriate to say that an intervention that shows minimal change but is well liked by participants is socially valid. For example, an exercise program may emphasize socializing while exercising (e.g., walking in small groups and talking). Attendance at sessions (one measure of outcome) may be very high. On a 7-point scale of "liking the program," the intervention may receive a mean rating of 6.5 ("like very much"). However, posttests may indicate the sessions resulted in a mean caloric expenditure by each participant of only 130 calories per session, and mean VO_2 max scores (maximum oxygen uptake; the definitive measure of cardiovascular fitness; Cooper, 1988) did not change. In this case, participants would remain in the "poor" fitness category despite their adherence to this well-liked program. Given appropriate effectiveness criteria, this would not be a socially valid program at any level of outcome analysis (e.g., Blair et al., 1989; Paffenbarger et al., 1986), except according to the sort of consumer satisfaction analyses recommended by Wolf (1978).

In the remainder of this paper, we illustrate various aspects of a broader definition of social validity by reviewing one of our ongoing interven-

tion research and development projects. The steps and problems involved in meeting various aspects of social validity are discussed.

CASE STUDY OF A SOCIALLY
VALID INTERVENTION

HIV Prevention with Adolescents

The purpose of the Family/Media AIDS Prevention Project (Winett *et al.*, 1991) is to design, develop, and evaluate (for efficacy and effectiveness) an HIV prevention videotape program for use in the home by parents and young adolescents (12 to 14 years of age). Prior to the start of the project, informal formative research by the first author had suggested (a) the home was an acceptable intervention site (*i.e.*, "place"); (b) there were few, if any, HIV education programs for parents and their teens (*i.e.*, "positioning"); (c) parents wanted a program to help them talk with their teens about topics such as sexual behavior; (d) a video program was seen as an ideal modality (*i.e.*, "product" and "promotion" considerations); and (e) a home-based program may reduce the time and effort required for learning and practice by parents and teens (*i.e.*, "price").

A primary objective of the program is to provide parents and teens with up-to-date information about HIV infection, other sexually transmitted diseases (STDs), and risk behaviors (alcohol and drug use) that are often antecedents to teens' engagement in nonprotected sexual activity. This objective pertains to knowledge gains. The behavioral objectives of the program are to teach parents and teens how to use problem-solving skills during family interactions and to teach teens how to use problem-solving and assertiveness skills in both everyday and high-risk situations. The use of the skills could eventually help teens avoid or manage other high-risk situations (*e.g.*, alcohol use at unchaperoned parties) more safely.

Another more general objective relates to beliefs and entails conveying the major conceptualization of the video program. The conceptualization is that parents can markedly reduce teenagers' risk of contracting HIV and other STDs, as well as decrease

the probability of other problems (*e.g.*, drug use), by discussing these topics explicitly, by motivating teens to stay involved in constructive activities (school, community groups, sports), and by learning and using specific skills offered in the program. The ultimate objective of the project involves the reduction of the incidence of HIV for teens who viewed the program with their parents and practiced skills in high-risk situations.

Background and Epidemiological Data

Adolescents are at risk for HIV because (a) most teens are sexually active by their middle to late teens, (b) only a small percentage of teens regularly use adequate protection during sexual activity, (c) many teen experiment sexually with multiple partners, (d) the rates of some other STDs (*e.g.*, syphilis) have increased for teens during the last decade, and (e) experimentation with alcohol and drugs increases the probability of engagement in other risk behaviors (*cf.* Miller, Turner, & Moses, 1990).

Present data on persons who have AIDS indicate that less than 1% of these cases are adolescents (CDC, 1991). However, these same data show that about 20% of these cases are persons in their early 20s (CDC, 1991). Because the incubation period from first infection to AIDS diagnosis can be 8 to 10 years, it is likely that the AIDS cases in this age range contracted HIV as adolescents (Miller *et al.*, 1990).

The current state of the AIDS epidemic is, perhaps, best represented by persons who are currently HIV positive but do not have AIDS and those persons at high risk for contracting HIV. A recent National Academy of Science report (Miller *et al.*, 1990) stated that HIV infection is no longer restricted geographically or demographically and the adolescent population has been "seeded" with HIV. Thus, focusing on adolescents in the general population seems verifiably important. Note, however, that our project focus is on all families with young teens and not just on those families who demographically or geographically are at greater risk for HIV. Although exact risk rates cannot be constructed for adolescents, data with other groups suggest that risk appreciably increases with alcohol

and drug involvement and multiple partners (Turner et al., 1990). These risk factors are widespread across segments of the teen population (Miller et al., 1990). Additionally, the risk of a teen contracting HIV from a sexual encounter increases as the rate of HIV increases in the teen population.

The choice of intervening with the family at home was further based on the limitations in explicitness of content in many school-based programs covering the same or similar topics, the possibility of parent and teen discussion of information, and parental guidance of teen practice and use of skills. In other words, parental involvement was seen as critical for environmental support of teen behavior change; in a sense, we have developed a multilevel intervention program (i.e., teen, parent, family).

Younger teens have been the focus of our preventive intervention program because most are not yet sexually active. The video modality was chosen because video programs can graphically illustrate major points (e.g., the appearance of a person in the terminal stages of AIDS) and demonstrate use of particular skills within a variety of specific situations. Additionally, a videotape program can be readily disseminated, because about 75% of U.S. homes have videocassette recorders. Clearly, a more preferred modality for skill development would be an interactive video program. However, in the 1990s relatively few homes have videodisc players, thus limiting the dissemination potential.

Social Marketing Research and Early Efficacy Tests

An extensive series of social marketing steps was undertaken. The purposes of these steps were to (a) develop methods for recruiting participants into the project, (b) reliably measure important cognitive (knowledge, beliefs) and behavioral outcome variables (i.e., specific skills taught, use of skills in everyday situations), (c) develop content acceptable to both parents and teens, and (d) determine whether the content could result in appropriate initial cognitive and behavior change. This last step is considered efficacy testing.

Table 1 reviews our social marketing steps and the initial efficacy testing we have conducted. Con-

tent experimentation also included format and formative feature changes (Moore, 1990). The steps illustrate the progression of content and measurement development involved in the project. Obviously, such development can only be accomplished with interactions among families representative of our target group (i.e., one- or two-parent families, lower-middle to upper-middle income, at least 1 boy or girl, 12 to 14 years old). However, all skill training sequences and exercises were based on procedures and formats from social cognitive theory (Bandura, 1986).

Steps 8 through 12 entailed using a number of formative and pilot research steps and initial efficacy tests in combination, resulting in totally new program content and refined measures. The new content appears to have increased the program's appeal (based on parent and teen ratings) for parents and teens. Note, however, Step 12 in Table 1. We found the greater appeal of a revised teen tape did not result in knowledge or skill gains unless there was active parental involvement during program viewing. Thus, by traditional definitions of social validity (Wolf, 1978), the new teen tape had high social validity (or consumer acceptance). By the broader definition of social validity proposed herein, however, the revised teen tape had no social validity when the teen viewed the tape alone, because the knowledge and skill gains were minimal. Not surprisingly, we are revising our video program once again and strongly emphasize parental involvement when viewing the videotape, and parental supervision of skill acquisition and practice (i.e., the original conception of the program).

However, even with objective evidence of knowledge and skill acquisition, for the video program to be considered effective (in contrast to the initial efficacy tests), it must eventually be shown that (a) parents and/or teens seek out the program from usual outlets (e.g., video stores) and gain appropriate knowledge and skills as a result, (b) parent and teen knowledge and skills developed from watching the video program together are maintained, (c) teens actually use the skills in appropriate situations, and (d) the incidence of risk behaviors and HIV infection and STDs are reduced on a large

Table 1
Social Marketing Research and Initial Efficacy Tests for the Family/Media AIDS Prevention Project

Step	Purpose	Procedure/Results
1. Family recruitment	Find effective way to recruit families into the project.	Compared recruitment through PTA and church groups, neighborhood canvasses, and through family physicians. PTA and church groups had many obstacles to recruitment. Neighborhood canvasses were inefficient. A three-step recruitment procedure with family physicians reliably recruited 50% of potential participants into the project. Used for the remainder of the project.
2. Family interviews	Gather information about parent/teen knowledge and conversations about target topics, teen engagement in risk behaviors, parent and teen ability to discuss target topics explicitly when given structured guidance.	In-depth (2.5 hr) home interviews with 13 families. Used different forms of knowledge test. Teen/parent general knowledge of topics was high; specific knowledge was low. One knowledge test developed and used in rest of the project. Teens and parents engage in frequent conversations with each other but have very few explicit conversations about target topics. Conversation style partly modifiable with in-home interviewer instruction and guidance.
3. Family skills/conversation assessment	Assess family conversations (topics, duration, frequency) teen/parent knowledge, teen/parent use of problem-solving skills, parent/teen belief concerning teen vulnerability to HIV and STDs, acceptability of potential content for the program.	In-depth (2.5 hr) home assessments conducted with eight families. Paper and pencil scales for conversations, knowledge, problem solving, and beliefs were used. Parents and teens role-played conversations and problem solving. Received brief feedback from assessors. Data confirm that conversations were frequent, but explicit conversations on target topics were infrequent. Specific HIV/STDs knowledge is low for parents and teens. Teen's frequency of risk behaviors was low. Parent/teens rate risk of HIV/STDs low for their teen or self. Conversations can be more explicit and problem-solving steps used with brief guidance and feedback from assessors. Problem-solving skills can be rated globally.
4. Family skills/conversation assessment	The same purposes as the prior step except paper and pencil forms were modified. The viability of self-monitoring of conversations was assessed, and the effectiveness of instruction and feedback on problem solving was assessed.	In-depth (2.5 hr) home assessments conducted with eight families. Knowledge, belief outcomes similar to Step 3. Parent and teen logs were discrepant. Feedback did not improve problem-solving ability.
5. Family skills/conversation assessment	The same purposes as the prior step except the logs were simplified and demonstrations of problem solving were shown to families in an 8-min (facsimile) video program. A role-play sequence to assess problem-solving skills was developed.	In-depth (2.5 hr) assessments done with four families. Problem-solving skills increased slightly with video instruction. The role-play sequences were too long and nonstandardized. The 5-point scale used in assessment only marginally discriminated skills. The log format was still complicated for most parents and teens.
6. Facsimile test, family skills training/skill assessment/conversation assessment	To assess if a brief video program alone could teach parents and teens problem-solving skills (compare video program with and without	Somewhat briefer (2 hr) assessments were done with 10 families. The brief video program could teach parents and teens problem-solving skills if the role-plays to assess the skills used

Table 1
(Continued)

Step	Purpose	Procedure/Results
	personal instruction), to compare different role-play assessment formats and conversation assessment procedures.	situations salient to families. Written communication logs were still not acceptable to parents and teens. A weekly call to parents and teens provided data on conversation in a more convenient way.
7. Focus group with teens 15 to 17 years old	To assess types of risk situations, degree of sexual activity, and acceptable coping skills for teens 1 to 3 years older than target teens.	Separate girls' ($n = 5$) and boys' ($n = 4$) focus groups were conducted. By 16, most teens were sexually active. High degree of alcohol use reported. Specific types of acceptable responses to risk situations were noted.
8. Pilot study, overall program (prototype test)	To assess the overall acceptability of an extended home video program for parents/teens, the conveying of information, and the teaching of assertion and problem-solving skills with the parent in the role of trainer.	A 90-min video program was produced and given to 23 families in a pretest, posttest, treatment-only study. The program used four separate tapes for information, training, practice, and "future" situations. The format was primarily didactic. Assertion and problem-solving were depicted in seven steps with some technical terms used. Emphasis placed on frequent teen/parent conversations and involvement of teens in positive activities. Skill assessments used role plays with standardized procedures, but not standardized scenes. The results showed knowledge increases for parents and teens, but minimal changes in skills or conversations.
9. Family interviews/focus groups	To assess strengths and weaknesses of the video program using feedback from families involved in Step 8.	The purpose of the program and a home-based video was highly endorsed. A less didactic, simplified program was recommended by parents and teens. The simplification of steps and terms was requested. A role of "coach" for the parent (not trainer) was recommended. More explicit content and more focus on skills was recommended.
10. Efficacy test, overall program	To present a completely redone, extended video program for parents and teens and assess efficacy by conveying information, teaching assertion and problem-solving skills with the parent in the role of coach.	A 120-min video program was produced. The program used four separate videotapes and included a dramatization of a typical family. Forty-four families were randomly assigned to experimental (video) or control (no video) conditions in a study with pretest, posttest, and followup phases; during follow-up, controls received the video program and were then reassessed. Assessments were done by two-person assessment teams blind to a family's condition. Role plays were standardized for assessment via an assessment video. Role plays were assessed at .90 reliability. Large knowledge gains were made by parents and teens, and large gains were seen in family problem-solving skills. Smaller but statistically significant gains were seen in teen assertiveness and problem solving. Video program was viewed only once or twice, and skill practice was limited. However, knowledge and skill gains were maintained at a 6-month follow-up, and parent and teen

Table 1
(Continued)

Step	Purpose	Procedure/Results
11. Family focus groups	To assess strengths and weaknesses of the video program using feedback from families in Step 10.	knowledge and skill gain were replicated when the control group viewed the program. The purpose of the program and a home-based video was highly endorsed. The use of the family dramatization was endorsed. Parents and teens requested content and skill practice steps that involved more explicit and difficult situations. The teens requested more teen characters and music. The connections between HIV, STDs, and present parent/teen relationships, teen skills, and activities needed to be clearer.
12. Efficacy test, teen program	To assess efficacy of a further simplified teen assertiveness and problem-solving video, skill training using more explicit and difficult situations and more specific connections between teen/parent relations, teen skills, teen activities, and HIV and STDs.	The parts of the video program teaching teen assertion and problem solving were completely reformatted and redone. The content reflected prior teen and parent input. A young adult narrator was used to introduce and develop rationale for skills, staying involved in positive activities, and communicating with parents. Thirty families were involved in a pretest, post-test, experimental, control condition test of the revised training video only. For some families, parental involvement in viewing was stressed. In other families, the video program was called a "teen only" program. Although teens rated this program more "likeable" than the last one, only with parental involvement were there knowledge and skill gains.

scale as a result of the program. These goals are ambitious, but they must also be tempered because this program is a primary prevention intervention. Because parents and teens use the video program at home with all the usual distractions and with no outside professional intervention, the degree and consistency of behavior change (i.e., appropriate use of skills in high-risk situations) will probably be modest. That is, outcomes at the individual level may appear marginal. However, the impact of even modest changes will be magnified at the population level if this video program or similar ones are widely available and viewed on a large scale by the general public.

CONCLUSION

This paper began with a broad definition of a socially valid, behavior-change intervention pro-

gram. To be considered socially valid, an intervention must be directed to problems of verifiable importance, the intervention must be accepted and used appropriately by designated target groups, and the intervention must have sufficient behavioral impact to reduce substantially the probability of the problem's occurrence in the target populations. We have emphasized that socially valid intervention programs are founded on epidemiological evidence, are designed by the use of conceptual frameworks through an interactive series of social marketing steps, and pass the tests of both efficacy and effectiveness studies. The various steps and processes involved in this broader definition of social validity were illustrated in part by a family-based HIV prevention program.

Meeting these criteria for a socially valid intervention program can be complicated, time-consuming, and often expensive. The overall approach is

far more extensive than Wolf's (1978) original conception of social validity, which was based only on the subjective evaluation of a program by its users. We have stressed in this paper that although the evaluations of program users are important and influence the design and refinement of an intervention program, such evaluations are only one component of a comprehensive analysis of social validity.

Granting that this broader approach to social validity is complicated, time-consuming, and expensive, what is the promise of the overall approach? The promise is that our behavior-change techniques and intervention programs will not only be well received by the public, but will truly make a verifiable difference in a nation's health and quality of life. Then we can call our preventive efforts socially valid.

REFERENCES

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Blair, S. N., Kohl, H. W., Paffenbarger, R. S., Clark, D. G., Cooper, K. H., & Gibbons, L. W. (1989). Physical fitness and all-cause mortality. *Journal of the American Medical Association*, **262**, 2395-2401.
- Bloom, B. (1988). *Health psychology: A psychological perspective*. Englewood Cliffs, NJ: Prentice-Hall.
- Centers for Disease Control. (1991, February). *HIV/AIDS surveillance data*. Washington, DC: National Center for Health Statistics.
- Committee on Diet and Health. (1989). *Diet and health: Implications for reducing chronic disease risk*. Washington, DC: National Academy of Science.
- Cooper, K. H. (1988). *Controlling cholesterol*. New York: Bantam Books.
- Farquhar, J. W., Fortmann, S. P., Flora, J. A., Taylor, C. B., Haskell, W. L., Williams, P. T., Maccoby, N., & Wood, P. D. (1990). Effects of community-wide education on cardiovascular risk factors. *Journal of the American Medical Association*, **264**, 359-365.
- Fawcett, S. B. (1990). *Balancing rigor and relevance: Some standards for behavioral research and community action*. Unpublished manuscript. University of Kansas, Lawrence, KS.
- Fawcett, S. B., Seekins, T., Whang, P., & Muiu, S. (1984). Creating and using social technologies for community empowerment. In J. Rappaport, C. Swift, & R. Hess (Eds.), *Studies in empowerment: Steps toward understanding and action*. *Prevention in Human Services*, **3**, 145-171.
- Fisher, M. (Ed.). (1989). *Guide to clinical preventive services: Report of the U.S. Preventive Services Task Force*. Baltimore, MD: Williams & Wilkins.
- Geller, E. S. (1989). Applied behavior analysis and social marketing: An integration for environmental preservation. *Journal of Social Issues*, **45**, 17-36.
- Geller, E. S. (1991, February). Preventing trauma from vehicle crashes with behavioral community psychology. *Behavior Therapist*, pp. 33-35.
- Greenwald, P., Sondik, E., & Lynch, B. S. (1986). Diet and chemoprevention in NCI's research strategy to achieve national cancer control objectives. In L. Breslow, J. E. Fielding, & L. B. Lave (Eds.), *Annual review of public health* (Vol. 7, pp. 267-292). Palo Alto, CA: Annual Reviews, Inc.
- Harris, S. S., Caspersen, C. J., DeFries, G. H., & Estes, E. H. (1989). Physical activity counseling for healthy adults as a primary preventive intervention in the clinical setting. *Journal of the American Medical Association*, **261**, 3590-3598.
- Henderson, M. M., Kushi, L. H., Thompson, D. J., Gorbach, S. L., Clifford, C. K., Insull, W., Moskowitz, M., & Thompson, R. S. (1990). Feasibility of a randomized trial of a low-fat diet for the prevention of breast cancer: Dietary compliance in the women's health trial Vanguard Study. *Preventive Medicine*, **19**, 115-133.
- Horwitz, S. D. (1989). *Let them call me rebel: Saul Alinsky, his life and legacy*. New York: Knopf.
- Hovell, M. F., Sallis, J. F., Hofstetter, C. R., Spry, V. M., Faucher, P., & Caspersen, C. J. (1989). Identifying correlates of walking for exercise: An epidemiologic prerequisite for physical activity promotion. *Preventive Medicine*, **18**, 856-866.
- Hursting, S. D., Thomquist, J., & Henderson, M. M. (1990). Types of dietary fat and the incidence of cancer at five sites. *Preventive Medicine*, **19**, 242-254.
- Jeffrey, R. W. (1989). Risk behavior and health: Contrasting individual and population perspectives. *American Psychologist*, **44**, 1194-1202.
- Keleman, M. H., Efron, M. B., Valenti, S. A., & Stewart, K. H. (1990). Exercise training combined with anti-hypertensive drug therapy: Effects on lipids, blood pressure, and left ventricular mass. *Journal of the American Medical Association*, **263**, 2766-2771.
- Koop, C. E. (1988). *The Surgeon General's report on nutrition and health*. Washington, DC: U.S. Department of Health and Human Services.
- Kotler, P., & Roberts, E. L. (1989). *Social marketing: Strategies for changing public behavior*. New York: Free Press.
- Kuller, L., Meilahn, E., Townsend, M., & Weinberg, G. (1982). Control of cigarette smoking from a medical perspective. In L. Breslow, J. E. Fielding, & L. B. Lave (Eds.), *Annual review of public health* (Vol. 3, pp. 153-178). Palo Alto, CA: Annual Reviews, Inc.
- Kwiterovich, P. (1989). *Beyond cholesterol: The Johns Hopkins complete guide for avoiding heart disease*. Baltimore, MD: Johns Hopkins University Press.
- Manoff, R. K. (1985). *Social marketing: Imperative for public health*. New York: Praeger.
- Miller, H. G., Turner, C. F., & Moses, L. E. (1990). *AIDS:*

- The second decade.* Washington, DC: National Academy Press.
- Moore, J. F. (1990, August). Instructional design technology and health behavior change. Paper in R. A. Winett, *Dissemination of health behavior change information in the mass media*. Symposium presented at the annual meeting of the American Psychological Association, Boston.
- Paffenbarger, R. S., Hyde, P. T., Wing, A. L., & Hsieh, C. (1986). Physical activity, all-cause mortality, and longevity of college alumni. *New England Journal of Medicine*, **314**, 605–613.
- Palinkas, L. A. (1985). Techniques of psychosocial epidemiology. In P. Karoly (Ed.), *Measurement strategies in health psychology* (pp. 49–113). New York: Wiley.
- Prochaska, J. O., & DiClemente, C. C. (1983). Stage process of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology*, **51**, 390–395.
- Robertson, L. S. (1986). Behavioral and environmental interventions for reducing motor vehicle trauma. In L. Breslow, J. E. Fielding, & L. B. Lave (Eds.), *Annual review of public health* (Vol. 7, pp. 13–34). Palo Alto, CA: Annual Reviews, Inc.
- Rogers, E. M. (1983). *Diffusion of innovation* (3rd ed.). New York: Free Press.
- Stolz, S. B. (1981). Adoption of innovation from applied behavioral research: "Does anybody care?" *Journal of Applied Behavior Analysis*, **14**, 491–505.
- Trevisan, M., Krogh, V., Freudenheim, J. L. Blake, A., Muti, P., Panico, S., Farinaro, E., Mancini, M., Menotti, A., Ricci, G., & The Research Group of the Italian National Research Council. (1990). Diet and coronary heart disease risk factors in a population with varied intake. *Preventive Medicine*, **19**, 231–245.
- Turner, C. F., Miller, H. G., & Moses, L. E. (1989). *AIDS: Sexual behavior and intravenous drug use*. Washington, DC: National Academy Press.
- Wandersman, A. (1981). A framework of participation in community organizations. *Journal of Applied Behavioral Science*, **17**, 27–58.
- Winett, R. A. (1986). *Information and behavior: Systems of influence*. Hillsdale, NJ: Erlbaum.
- Winett, R. A., Altman, D. G., & King, A. C. (1990). Conceptual and strategic foundations for effective media campaigns for preventing the spread of HIV infection. *Evaluation and Program Planning*, **13**, 91–104.
- Winett, R. A., King, A. C., & Altman, D. G. (1989). *Health psychology and public health: An integrative approach*. Elmsford, NY: Pergamon Press.
- Winett, R. A., Moore, J. F., Anderson, E. S., Sikkema, K. J., Hook, R., Webster, D. A., Taylor, C. D., & Dalton, J. (1991). *Family/media approach to AIDS prevention. Interim report*. Blacksburg, VA: Virginia Polytechnic Institute & State University.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement, or how applied behavior analysis is finding it heart. *Journal of Applied Behavior Analysis*, **11**, 203–214.

Received August 10, 1990

Initial editorial decision January 15, 1991

Revisions received February 11, 1991; February 25, 1991; March 1, 1991

Final acceptance March 7, 1991

Action Editor, E. Scott Geller