

The Best of Practices, the Worst of Practices: The Making of Science-Based Primary Prevention Programs

Dennis M. Gorman, Ph.D., M.Sc.

Do school-based drug and alcohol prevention programs have any effect on subsequent use of these substances by students? Twenty years ago the answer would have been a resounding “no.” Many programs were in fact found to have negative effects, leading a 1973 government report to recommend a moratorium on school-based drug prevention (1). All this changed, however, in the mid-1980s with the “discovery” of the social influence approach that purported to teach children the skills necessary to resist social pressure to use drugs. Read almost any literature review in this area, and the story unfolds as follows.

In the 1960s and 1970s we relied on strategies—first giving factual information and then trying to make children feel good about themselves (“affective” education)—that research showed to be ineffective (2,3). In the mid-1980s, researchers began to develop and test programs that were based on sound psychological theories, such as social learning theory and problem behavior theory. Evaluations of these programs showed that they were successful in preventing a wide range of undesirable behaviors, including alcohol and drug use. Thus, unlike previous approaches that were not theoretically

grounded or supported by empirical findings, the new approach had the distinction of being “science-based.”

The trickle-down effect from the research world to the frontline practitioner was at best a dribble; there was little institutional diffusion of social influence programs during the early 1990s (4). This state of affairs might have prevailed had illicit drug use not increased in the mid-1990s, leading politicians and others to question what we were getting for the money spent on prevention efforts. In response, federal agencies began to demand that recipients of funds use only programs that were supported by scientific evaluation research. As a result, almost every federal agency with responsibility for drug prevention has produced a “best practice” or “science-based” list of approved prevention programs during the past five years (5–7).

Most of the school-based programs that appear on these best-practice lists are variants of the social influence approach. Unfortunately, it is difficult to use the term “best practice” to describe the evaluation research that is conducted in relation to many of these interventions. In fact, much of what goes on in the analysis of these school-based prevention programs is simply not consistent with the type of rigorous hypothesis testing that one associates with the term “science” and that has been a mainstay of evaluation research for the past 25 years (8). I have described many of these practices—such as multiple subgroup analysis, post hoc sample refinement, and use of points in time other than the study baseline to calculate attrition rates—

in a number of recent publications (9–12). In this column I briefly describe two other common practices, using as examples three of the most widely advocated prevention programs—the Seattle Social Development Project (SSDP), the Life Skills Training (LST) program, and the ATLAS program.

The adjustable outcome

The adjustable outcome can take two forms, one more subtle than the other. In its most obvious form, the adjustable outcome involves a total change in outcome over the course of the evaluation. For example, the SSDP, developed by David Hawkins and colleagues at the University of Washington, has been promoted as a delinquency and drug use prevention program for the best part of 20 years (13), but recent evaluations show that it has very little effect on such behaviors (12,14,15). However, because these evaluations have produced some positive effects on some health-related sexual practices, such as condom use, among some of the intervention group, the SSDP is now being heralded as a sex prevention program (15,16).

In its more subtle form, the adjustable outcome involves a change in the way the variable is constructed from study to study rather than a total change in target outcomes. For example, a recent critique of a longitudinal evaluation of Gilbert Botvin’s LST program (17) observes that there was a shift from the use of continuous outcome measures—for example, a 9-point scale of marijuana use—in the initial report from the study (18) to less sensitive dichotomous yes-or-no

Dr. Gorman is affiliated with the Department of Epidemiology and Biostatistics, School of Public Health, Texas A&M University System Health Science Center, 3000 Briarcrest Drive, Suite 310, Wells Fargo Building, Bryan, Texas 77802 (e-mail, gorman@srph.tamushsc.edu). Sally L. Satel, M.D., is editor of this column.

measures in a later paper published in *JAMA* (19). This begs the question as to whether the results of the trial are measurement dependent. Specifically, would the positive program effects reported for a subsample of the study population in the *JAMA* paper have emerged from an analysis based on the continuous measures used in the initial report?

Moreover, when one looks across the entire body of LST evaluations, there are other similar instances in which the manner in which outcome variables are constructed changes from study to study, and even from report to report. For example, a 1992 study of the effects of the program on smoking among students from 47 New York City schools (20) presented outcome data in terms of three dichotomous measures (use during the past month, the past week, and the current day) as well as an 11-point quantity scale, whereas a 2001 report from a study of students in 29 schools in the city (21) used just two continuous scales (a 9-point frequency and an 11-point quantity scale). Furthermore, a 2003 report from the latter study presented smoking outcomes in terms of a composite measure that combined the mean score on the 9-point frequency scale with those on an 8-point quantity scale (22). Again, one is left wondering whether the consistency in effects claimed for the program would exist had a consistent method been used in operationalizing the behavioral outcomes.

The inflatable p value and the wishful one-tailed test

Although nothing is carved in stone about the level at which statistical significance is set, the traditional level is .05—that is, accepting a 5 percent probability that one's findings are due to chance. Using a p value of .1 obviously doubles the likelihood of finding a statistically significant effect, as does the use of one-tailed significance tests since the critical region is then shifted to one end of the distribution. Statisticians typically recommend that researchers limit use of one-tailed significance tests to situations in which there is a very strong prior hypothesis (23). In the case of evaluations of prevention programs, their

use would be limited to instances in which previous research led one to expect only positive, not negative, effects on outcomes in such situations.

The 1995 analysis of the LST program (19) used one-tailed tests in examining effects on alcohol and marijuana use, even though a previous published evaluation of the program (24) showed that it had virtually no effect on participants' marijuana use and actually had some negative effects on use of alcohol. Similarly, a recent evaluation of the ATLAS program (25) used such tests in assessing the effects of the program on steroid use, even though an earlier report from the study (26) indicated that the program had no effect on this behavioral outcome (11). Thus there was no empirical basis for using one-tailed significance tests with the behavioral outcomes in either study.

Thinking scientifically and thinking emotionally

When I raise such concerns about the evaluation practices employed in school-based prevention research at conferences or among colleagues, I frequently get three responses—none of which has too much to do with a science-based approach to prevention. The first response is usually along the lines of “Who are you to criticize these programs and the accompanying research when experts have declared them effective?” Such an attitude is decidedly antiscientific, because it fundamentally rejects a basic tenet of critical thinking—namely, that one judge a thought or idea on the basis of its content, not by the person of the thinker (27).

The second response to my critique takes the form of “You shouldn't criticize these programs unless you have some alternative intervention to recommend.” If the premise of this argument were accepted, then intervention studies would be subject to scrutiny only from those who have developed an alternative that has been shown to be effective; all others would be prohibited from commenting. Clearly, though, one's ability to assess the methodological soundness of an evaluation is in no way dependent on whether one has developed an alternative intervention program.

Thus such an argument is entirely without any internal logic and hence is decidedly unscientific.

The final common response to my criticisms of evaluations of school-based prevention programs goes something like “Well, yes, we want prevention to be scientific, but the criteria you invoke are simply too strict. As a ‘new science,’ prevention should be able to bend the methodological rules a bit, especially in the interests of a good cause and with a program that we just know deep down must do some good.” The problem with this argument is that these methodological rules and procedures have a purpose—namely, to isolate the effects of one's program from other influences that can bring about the type of behavioral change we desire—and when they are bent too much they no longer fulfill this purpose. Also, the rule bending tends not to be very evenly applied, and one is left wondering how prevention strategies that researchers have evaluated in the past and found to be ineffective—such as affective education and the Drug Abuse Resistance Education (DARE) program—would have fared had they been assessed with one-tailed significance tests, multiple subgroup analysis, and the like.

James L. Nolan, Jr., has observed that today's therapeutic state must balance its concerns about self-actualization and self-fulfillment with a utilitarian concern about effectiveness, efficiency, and cost. For the most part, the two coexist and even complement one another. However, when they come into conflict it is intuition and emotion that prevail over logic and reason (28). Science-based school prevention programs are a prime example of this balancing act, and the response to critiques of evaluations of these programs demonstrates the triumph of the therapeutic perspective over the scientific. ♦

Acknowledgment

This work was funded by a grant from the Smith Richardson Foundation, Inc.

References

1. Drug Use in America: Problem in Perspective. Washington, DC, National Commission on Marijuana and Drug Abuse, 1973

2. Botvin GJ: Substance abuse prevention research: recent developments and future directions. *Journal of School Health* 56:369–374, 1986
3. Ellickson PL: Schools, in *Handbook of Drug Abuse Prevention*. Edited by Coombs RH, Ziedonis D. Boston, Allyn & Bacon, 1995
4. Rohrbach LA, D'Onofrio CN, Backer TE, et al: Diffusion of school-based substance abuse prevention programs. *American Behavioral Scientist* 39:919–934, 1996
5. Preventing Drug Use Among Children and Adolescents: A Research-Based Guide. Rockville, Md, National Institute on Drug Abuse, 1997
6. 2001 Annual Report of Science-Based Prevention Programs. Rockville, Md, Center for Substance Abuse Prevention, 2001
7. US Department of Education: Safe, Disciplined, and Drug-Free Schools Expert Panel: Exemplary Programs, 2001. Available at www.ed.gov/offices/oeri/orad/kad/expert_panel/2001exemplary_sdfs.html
8. Cook TD, Campbell DT: *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago, Rand McNally, 1979
9. Gorman DM: The irrelevance of evidence in the development of school-based drug prevention policy, 1986–1996. *Evaluation Review* 22:118–146, 1998
10. Gorman DM: The “science” of drug and alcohol prevention: the case of the randomized trial of the Life Skills Training (LST) Program. *International Journal of Drug Policy* 13:21–26, 2002
11. Gorman DM: Defining and operationalizing “research-based” prevention: a critique (with case studies) of the US Department of Education’s Safe, Disciplined, and Drug-Free Schools Exemplary Programs. *Evaluation and Program Planning* 25:295–302, 2002
12. Gorman DM: Overstating the behavioral effects of the Seattle Social Development Project. *Archives of Pediatric and Adolescent Medicine* 156:155–156, 2002
13. Hawkins JD, Weis JG: The social development model: an integrated approach to delinquency prevention. *Journal of Primary Prevention* 6:73–97, 1985
14. Hawkins JD, Catalano RF, Kosterman R, et al: Preventing adolescent health risk behaviors by strengthening protection during childhood. *Archives of Pediatric and Adolescent Medicine* 153:226–234, 1999
15. Lonczak HS, Abbott RD, Hawkins JD, et al: Effects of the Seattle Social Development Project on sexual behavior, pregnancy, birth, and sexually transmitted disease outcomes by age 21 years. *Archives of Pediatric and Adolescent Medicine* 156:438–447, 2002
16. Ethier K, St Lawrence JS: The role of early, multilevel youth development programs in preventing health risk behavior in adolescents and young adults (editorial). *Archives of Pediatric and Adolescent Medicine* 156:429–430, 2002
17. Brown JH: Youth, drugs, and resilience education. *Journal of Drug Education* 31: 83–122, 2001
18. Botvin GJ, Baker E, Dusenbury L, et al: Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: results of a 3-year study. *Journal of Consulting and Clinical Psychology* 58:437–446, 1990
19. Botvin GJ, Baker E, Dusenbury L, et al: Long-term follow-up of a randomized drug abuse prevention trial in a white middle-class population. *JAMA* 273: 1106–1112, 1995
20. Botvin GJ, Dusenbury L, Baker E, et al: Smoking prevention among urban minority youth: assessing effects on outcome and mediating variables. *Health Psychology* 11:290–299, 1992
21. Botvin GJ, Griffin KW, Diaz T, et al: Drug abuse prevention among minority adolescents: posttest and one-year follow-up of school-based preventive intervention. *Prevention Science* 2:1–13, 2001
22. Griffith KW, Botvin GJ, Nichols TR, et al: Effectiveness of a universal drug abuse prevention approach for youth at high risk for substance use initiation. *Preventive Medicine* 36:1–7, 2003
23. Blalock HM: *Social Statistics*, 2nd ed rev. London, McGraw-Hill, 1979
24. Botvin GJ, Baker E, Filazzola AD, et al: A cognitive-behavioral approach to substance abuse prevention: one-year follow-up. *Addictive Behaviors* 15:47–63, 1990
25. Goldberg L, MacKinnon DP, Elliot DL, et al: The Adolescent Training and Learning to Avoid Steroids Program: preventing drug use and promoting health behaviors. *Archives of Pediatric and Adolescent Medicine* 154:332–338, 2000
26. Goldberg L, Elliot D, Clarke GN, et al: Effects of a multidimensional anabolic steroid prevention intervention: the Adolescent Training and Learning to Avoid Steroids (ATLAS) Program. *JAMA* 276: 1555–1562, 1996
27. Notturmo MA: *Science and the Open Society: the Future of Karl Popper’s Philosophy*. New York, CEU Press, 2000
28. Nolan Jr JL: *The Therapeutic State: Justifying Government at Century’s End*. New York, New York University Press, 1998

BEST PRACTICES

Continued from page 1080

♦ If a company is presenting potentially biased information—for example, comparison data that are favorable to its own product—is there an opportunity for other views and data to be presented?

♦ Is the opportunity for support equitably distributed among staff or programs within the facility?

In summary, we have presented a best-practices model, developed by a medical and professional staff, to sort through the complex issues of accepting resources from pharmaceutical companies within the culture of an academic mental health center. By taking a risk-benefit approach to the specific requests and opportunities and by asking relevant ethical and practical questions, our group was able to reach consensus decisions and provide an educational structure from which flexible guidelines are derived. This model provides a conceptual framework for approaching ethical decisions and is readily adaptable to other settings. Clearly the model will need to be evaluated over time and in other settings to assess any further

positive or negative impact on patient care, the educational program, and interprofessional relationships. ♦

References

1. Lichstein PR, Turner RC, O'Brien K: Impact of pharmaceutical company representatives on internal medicine residency programs: a survey of residency program directors. *Archives of Internal Medicine* 152: 1009–1013, 1992
2. Rosner F: Pharmaceutical industry support for continuing medical education programs: a review of current ethical guidelines. *Mount Sinai Journal of Medicine* 62: 427–430, 1995
3. Schwartz TL, Kuhles DJ, Wade M, et al: Newly admitted psychiatric patient prescriptions and pharmaceutical sales visits. *Annals of Clinical Psychiatry* 13:159–162, 2001
4. Wazana A: Physicians and the pharmaceutical industry: is a gift ever just a gift? *JAMA* 283:373–380, 2000
5. An unhealthy influence on doctors (editorial). *New York Times*, Sept 10, 2001
6. Vendantam S: Industry role in medical meeting decried. *Washington Post*, May 25, 2002
7. ABIM Foundation, ACT-ASIM Foundation, and European Federation of Internal Medicine: Medical professionalism in the new millennium: a physician charter. *Annals of Internal Medicine* 136:243–246