

Characteristics of Effective School-Based Substance Abuse Prevention

Denise C. Gottfredson^{1,3} and David B. Wilson^{1,2}

This study summarizes, using meta-analytic techniques, results from 94 studies of school-based prevention activities that examined alcohol or other drug use outcomes. It set out to determine what features of school-based substance abuse prevention programs are related to variability in the size of program effects. It asked (1) Which populations (e.g., high risk vs. general population) should be targeted for prevention services? (2) What is the best age or developmental stage for prevention programming? (3) Does program duration matter? and (4) Does the role of the person delivering the service (e.g., teacher, law enforcement officer, peer) matter? The results suggest that targeting middle school aged children and designing programs that can be delivered primarily by peer leaders will increase the effectiveness of school-based substance use prevention programs. The results also imply that such programs need not be lengthy. The evidence related to the targeting issue is sparse, but suggests that, at least for programs teaching social competency skills, targeting higher risk youths may yield stronger effects than targeting the general population. Suggestions for future research are offered.

KEY WORDS: substance abuse; prevention; school.

INTRODUCTION

Schools use a large number of different strategies to reduce or prevent youth substance use and other problem behaviors. A recent national study of school-based prevention (Gottfredson & Gottfredson, 2001) showed that the median number of different prevention activities currently underway in the typical school was 14. The number of such activities ranged from 0 in some schools to 66 in others. Schools also make regular use of a broad array of activities. All schools use rules and policies to deter substance possession and use or other problem behaviors. Nine of

10 schools provide information about tobacco, alcohol, other drugs, violence, accidents, health or mental health, or risky sexual behavior. Three quarters of schools provide prevention lessons in the form of curriculum instruction. Many schools employ interventions directed at students as well as at organizational or environmental arrangements. Curriculum and instruction, counseling, behavioral and recreation programs are among the former. Alterations in schedules, architectural features, and instructional or classroom management practices are among the latter. Just over half of America's schools have explicit security or surveillance programs. Little is known about the effectiveness of most of these school-based activities for reducing or preventing substance use.

Reviews of research on school-based substance abuse prevention have concluded that some forms of school-based prevention are effective and have identified a few characteristics related to effectiveness. The following section summarizes what is known about the effectiveness of school-based substance abuse prevention.

¹Department of Criminology and Criminal Justice, University of Maryland, College Park, Maryland.

²Present address: Administration of Justice Department, George Mason University, Manassas, Virginia.

³Correspondence should be directed to Denise C. Gottfredson, Department of Criminology and Criminal Justice, 2220D LeFrak Hall, University of Maryland, College Park, Maryland 20742; e-mail: dgottfredson@crim.umd.edu.

Previous Reviews of Substance Abuse Prevention⁴

In the last decade many reviews of substance abuse prevention have been published (Botvin, 1990; Botvin *et al.*, 1995; Dryfoos, 1990; Durlak, 1995; Ennett *et al.*, 1994; Gerstein & Green, 1993; Gorman, 1995; Gottfredson, 1997, 2001; Gottfredson *et al.*, 2002; Hansen, 1992; Hansen & O'Malley, 1996; Hawkins *et al.*, 1995; Institute of Medicine, 1994; Norman & Turner, 1993; Tobler, 1992; Tobler & Stratton, 1997). These reviews concluded that some forms of preventive activities have been demonstrated to reduce substance use. Botvin (1990) concluded that approaches that include resistance-skills training to teach students about social influences to engage in substance use and specific skills for effectively resisting these pressures alone or in combination with broad-based life-skills training reduce substance use.

Subsequent reviews generally concurred with Botvin's conclusions regarding the relative effectiveness of social skills approaches as opposed to information-only and affective approaches to classroom-based instructional programs, but they raised questions about the use of different modalities (other than classroom instruction), and suggested that different content (other than social skills training) might also be effective in classroom instruction. For example, Hansen (1992) and Gottfredson (1997) indicated that approaches aimed at changing normative beliefs about drug use are effective. These approaches often use survey results to correct misperceptions about the prevalence of use, engage youths in discussions to elicit their opinions about the appropriateness, and include testimonials from admired peers emphasizing that use is not acceptable. Instructional programs that incorporate these norm-setting activities have been shown to reduce use (Gottfredson *et al.*, 2002; Hansen, 1992), but noninstructional programs that employ these methods outside the classroom are also effective for reducing substance use (Hansen & Graham, 1991; Perry *et al.*, 1996).

A recent meta-analysis (Wilson *et al.*, 2001) compared effect sizes across 11 intervention categories of school-based prevention, including 7 types of individually focused and 4 types of environmentally focused interventions. The study concluded that environmentally focused interventions are generally

more effective for reducing alcohol and other drug (AOD) use than are individually focused interventions. But among the latter, cognitive-behaviorally and behaviorally based interventions, both with and without an instructional self-control or social competency component, are the most effective. On the other hand, counseling that is not based on a behavioral or cognitive-behavioral model, mentoring, tutoring and work study programs, and recreational programs were found to be ineffective. Unfortunately, the number of studies on which these conclusions of ineffectiveness were based is small. The report concluded that additional research on a wider variety of plausible strategies for substance abuse prevention is sorely needed. It also showed that although the type of school-based strategy accounted for significant variation in the effect sizes, substantial unexplained variance remained after partialling-out modality. Other factors must also contribute to program effectiveness.

Reviews have suggested that the delivery mechanism and methods, duration, and timing may be important moderator variables. Tobler and Stratton (1997), for example, reviewed studies of school-based drug prevention programs that were available to the entire student body and targeted grades 6–12. They concluded that “interactive” programs (e.g., those affording much opportunity for interaction among the adolescents) were more effective than “noninteractive” programs (e.g., didactic presentations). They also showed that program content categories (e.g., social influence, information only, affective) were correlated with mode of delivery (interactive vs. noninteractive), and suggested that some of the positive effect previously attributed to program content may in fact be due to the delivery method.

Reviews have also concluded that the role of the deliverer is important. Hansen (1992) suggested that the training and background of the leader and the fidelity of presentation may be more important than the content of the message. Tobler compared effect sizes for programs delivered by different types of leaders. Mental health professionals and counselors produced the largest effects, followed by peers and then teachers. Tobler (1992, pp. 20–21) concluded that the leader must be someone who is

competent in group process, who can enhance the interactional process and simultaneously focus and direct the group. Successful leaders have the ability to act as guides, as opposed to being dominant. They are able to tolerate ambivalence, and know when to remain silent to facilitate true dialogue. They are able to empower adolescents to make conscientious

⁴Review based on an earlier review of prevention programs that appeared in Committee on Data and Research for Policy on Illegal Drugs (2001).

decisions and to encourage freedom of choice and individual self determination.

Most likely, the content of the message and the characteristics of the leader interact to produce more or less effective programs, but this assertion has not been subjected to an empirical test.

Previous reviews have also been inconclusive regarding the optimal timing of delivery of substance abuse prevention. Because many of the risk factors for later substance use appear at an early age, it can be argued that prevention should start early. Many prevention efforts, however, are delivered during early adolescence, when social influence to use increases. Only a few studies have assessed effects on later substance use of attempts to alter risk factors that appear at an earlier developmental stage, but they have generally produced positive results on subsequent substance use outcomes. These studies of early prevention efforts include Lochman's "anger coping" program (Lochman, 1992), the Good Behavior Game (Dolan *et al.*, 1993; Kellam *et al.*, 1994; Kellam & Anthony, 1998), the Child Development Project (Battistich *et al.*, 1996), and the School Development Program (Comer, 1985; Cook *et al.*, 1998). Although these examples provide evidence that early prevention works to reduce substance use, we do not know how the magnitude of the effects of these activities compares with those of programs that target older youths.

The duration of programs also varies considerably from study to study. The accumulated wisdom in the prevention field (e.g., Durlak, 1995; Gottfredson, 1997) is that longer is better: Programs that provide "booster" sessions after the initial activity produce more lasting effects than those that do not. These conclusions are based largely on a handful of studies that have compared effects of an instructional program with and without a booster. Botvin *et al.* (1984), for example, show that the effects of peer-led Life Skills Training on self-reports of marijuana use in the past month taken 16 months after the initial pretest are not significantly different from zero (effect size = 0.22) for those students in the condition without the booster, but when additional lessons are provided in the following school year to reinforce the initial lessons, effects at 16 months after the pretest are significant and more than doubled in magnitude (effect size = 0.52). However, others have demonstrated both short-term (Eggert *et al.*, 1994) and long-term (e.g., Lochman, 1992) positive effects of "one shot" interventions of approximately the same dura-

tion as Botvin's initial program. Additional clarity is needed to understand the value added through longer programs.

Another important design feature about which little is known is the target population for substance abuse prevention activities. Much of the research on school-based prevention focuses on "universal" programs. This approach is based on the earlier research (e.g., Robins & Przybeck, 1985) demonstrating that early initiation of drug use leads to more frequent and problematic use later on. Universal programs target the general student population in an attempt to delay early initiation for all. An equally plausible approach to prevention is to target prevention activities as a subpopulation of youths identified as at especially high risk for developing serious drug problems. This approach is consistent with criminological research (e.g., Moffitt, 1993) that has demonstrated that a small subset of the general population is responsible for the majority of the crime that occurs, and that individuals who engage in criminal activity are highly likely to also engage in more frequent and problematic substance use. Huizinga and Jakob-Chien (1998), for example, show that between 37 and 51% (depending on the sample) of youths aged 13–17 who committed serious crimes also used drugs, compared with between 1 and 3% among nondelinquent youths. These criminological facts argue for an approach that targets youths with elevated likelihood for engaging in a variety of misbehavior (including substance use) for prevention services.

Reviews and meta-analyses of prevention effectiveness fail to differentiate among programs that target at-risk and universal populations. A handful of studies have compared the effectiveness of universal prevention activities for groups that differed according to their level of use at baseline. These studies (e.g., Bell *et al.*, 1993; Ellickson *et al.*, 1993; Ellickson & Bell, 1990; Hansen *et al.*, 1988) have suggested that universal prevention programs designed to prevent initiation in the general population may either increase use or have no effect on use for the most at-risk segment of the population. More specialized programming may be needed for the most at-risk segments of the population, but existing research provides little or no guidance on this issue.

This study is designed to increase knowledge about these features of school-based substance abuse prevention. It summarizes, using meta-analytic techniques, results from 94 studies of school-based prevention activities that examined alcohol or other drug use outcomes. These activities ranged from individual

counseling or behavior modification programs through efforts to change the way schools are managed. The study is intended to determine, on the basis of currently available studies, what features of school-based substance abuse prevention programs are related to variability in the size of program effects. It asks (1) Which populations (e.g., high risk vs. general population) should be targeted for prevention services? (2) What is the best age or developmental stage for prevention programming? (3) Does total program duration matter? and (4) Does the role of the person delivering the service (e.g., teacher, law enforcement officer, peer) matter? This meta-analysis differs from our previous work in that it provides a more detailed analysis of AOD outcomes and examines the above potential moderators of program effectiveness.

METHODS

Inclusion Criteria and Search for Relevant Studies

To be included in this report, a study had to meet the following criteria: (a) it evaluated an intervention, that is, a distinct program or procedure intended to reduce problem behaviors among children and youth; (b) the intervention was school-based, that is, the intervention was integrated into the school environment, such as classroom instruction or other classroom activities, a school-wide environmental change, or modification of teacher or school administrators behaviors and instructional practices; (c) it used a comparison group evaluation methodology, including nonequivalent comparison group research designs, and the comparison group was a no-treatment or minimal-treatment condition; and (d) it measured AOD use, excluding cigarette and smokeless tobacco use.⁵ Excluded from this synthesis were studies that did not report sufficient information to determine the direction of the effect for an AOD outcome. Studies were included even if they did not report sufficient information to directly compute an effect size, contingent on our ability to determine the sample size and whether the intervention or comparison was favored on at least one outcome of interest. In these cases, we imputed an effect size based on statistical significance, if reported. In some cases, effect sizes

were imputed for studies that only reported that an effect was statistically nonsignificant. The imputation was based on the sample size, direction of effect, and an assumed one-tailed p -value of .25. This generated a small effect size in the reported direction.

Potentially eligible studies for inclusion in this synthesis were identified through searches of bibliographic databases (e.g., Digital Dissertations Online, PsychLit, ERIC, and Sociological Abstracts) and through the references of recent reviews of prevention programs cited above. In some instances, the search of recent reviews resulted in the identification and inclusion of a number of unpublished studies. This list was augmented with additional studies already known to the authors. A total of 94 studies, representing 130 documents, met these criteria.

Effect sizes were based on the contrast between a treatment and comparison group condition. Some studies had multiple treatment–comparison group contrasts. In this meta-analysis, we included multiple treatment–comparison group contrasts from a single study when the contrasts represented unique interventions or when the program participants represented distinct groups of students. In some cases, these multiple treatment–comparison contrasts shared a comparison group. Although this introduced statistical dependencies in the data, this was balanced against the potential benefit of examining these programs in a more differentiated fashion. Furthermore, the statistical model used to analyze the effect sizes incorporated the information on these statistical dependencies (see Wilson *et al.*, 2001). In all, 136 contrasts were coded from the 94 studies included in the synthesis.

Coding of Study Characteristics

A code book similar to a survey form was developed to capture information regarding the specific nature of the intervention, characteristics of the student population, research methodology, measures of AOD use, and observed effects on these measures at all measurement points. Studies were coded by trained graduate students who meet weekly to discuss coding decisions. To improve reliability, all studies were coded by at least two coders and all coding discrepancies were discussed and resolved. A copy of the code book can be obtained from the authors.

Program Characteristics

The nature of the student population participating in the school-based prevention program was

⁵Studies using cigarette or smokeless tobacco as the only measure of substance use were excluded purely for practical reasons owing to the large number of such studies and the limited resources available for coding.

captured by a set of items addressing the age and grade range represented in the study. The coding protocol also captured whether the study represented the general school population or was restricted in some way to a high risk group, such as youths with a prior criminal history or reported drug use. Also coded were the variables indicating the type of individual(s) providing the program (e.g., teacher, police officer, peers) and the duration of the program. Duration was measured simply as the length of the program in months from the beginning until the end of the last session (including booster sessions). Summer months were subtracted. Although it would have been useful to know the total number of hours of contact time a program had with the students, this information was too rarely provided to be used analytically.

Characteristics of the Research

The soundness of the empirical evidence presented in each study was also measured so that these differences could be statistically controlled. This was accomplished through seven items in the coding protocol that assessed methodological aspects of the research study. These items addressed assignment to conditions (e.g., random assignment to conditions), unit-of-assignment (e.g., student, class, school), unit-of-analysis, use of control variables in analyses to adjust for initial group differences, rating of initial group similarity, attrition, and an overall 5-point evaluation of methodological quality.

Program Effects

The effectiveness of the program on each available outcome was coded using the standardized mean difference effect size, a measure of the difference between the program and comparison groups relative to the standard deviation of the measure employed. Whenever possible, the posttreatment or follow-up mean-difference was adjusted for any baseline mean difference on that measure. The standardization of the program and comparison difference by use of the effect size index allows for the direct comparison of effects across studies and outcomes. Furthermore, the standardized mean difference effect size can be computed from a wide variety of data configurations reported by the primary studies (Lipsey & Wilson, 2001).

Studies often reported on multiple drug use measures and may also have collected data at multiple

time points, postintervention. As is common in meta-analysis, all effects measuring AOD were averaged within a study.

In some cases, an effect size was estimated based on limited information, such as an imprecise probability level (e.g., $p < .05$) or a simple statement that the conditions were the same on the outcome. Ignoring these effects would lead to an upward bias in the mean effects. It is common practice in meta-analysis to estimate effect sizes in these situations, as was done here. Analyses were performed both with and without these estimated effects. No important differences were found. As such, only the analyses using the full data set are provided below.

RESULTS

This collection of studies covers a broad range of school-based prevention activities. The distribution of effects for delinquency, AOD use, truancy/drop out, and other problem behaviors has been discussed elsewhere (Wilson *et al.*, 2001). The focus of this paper is the relationship between the AOD use outcomes and characteristics of the population (i.e., high risk vs. general school population and age or development stage of the program participants), and selected characteristics of the intervention program (i.e., program duration and person delivering the service). Each of these relationships is explored below. An analysis of the robustness of the findings to methodological variability across studies is also presented.

Characteristics of the Student Sample

The majority of studies in this synthesis evaluated a universal prevention program, that is, one that was provided to the general student population. Only 11 of the 136 treatment-comparison contrasts targeted high-risk student samples. The random effects inverse variance mean effect size (see Lipsey & Wilson, 2001; Wilson *et al.*, 2001, for formulas) for the high-risk and general student population studies were roughly comparable (0.05 and 0.07, $p \leq .05$, respectively). This finding was counter to expectation because in our previous meta-analysis (Wilson *et al.*, 2001) that included studies of a wide range of problem behaviors and included a substantially larger number of effect sizes obtained from studies of high-risk students (61), more positive effects were found for programs targeting high-risk students. Several of the 11 studies with the

high-risk youth evaluated treatment modalities that were shown in our previous meta-analysis to be ineffective or minimally effective. Restricting the analysis to the cognitive-behaviorally oriented programs reveals a larger mean effect for the studies serving high-risk youths (effect size = 0.20, $n = 5$) than the general school population studies (effect size = 0.05, $n = 39$). This difference does not reach conventional levels of statistical significance owing to the small number of studies for the high-risk mean effect size, despite being a fourfold difference in effectiveness. Thus there is some evidence in support of the hypothesis that programs targeting high-risk youth are more effective. This evidence is weak, however, and needs further replication and elaboration. Furthermore, this relationship may be moderated by the type of intervention.

A second student characteristic of interest to this synthesis is the age or developmental stage of the students. Assessing the relationship between program effectiveness and age is complicated by the varying age spans of the study samples, ranging from studies with samples of a single school grade to studies with samples spanning 1st through 12th grade. Most studies, however, were restricted to either late elementary, middle or junior high school, or senior high school students. The mean effect sizes for each of these grade categories is presented in Table 1. The results suggest a slight advantage for school-based AOD prevention programs delivered to middle/junior high school students, although the difference is not statistically significant. The mean effect size is roughly double in size of that for studies of late elementary and senior high school students.

This comparison may be confounded, however, by the varying lengths of follow-up for the different studies. Studies of late elementary school programs, on average, measured outcomes 2-year postintervention, whereas evaluations of mid-

dle/junior high school programs had a follow-up period of 1-year, on average. Evaluations of senior high school programs tended to measure drug use 6-month to 1-year postintervention. It is possible that studies of elementary school students are disadvantaged in this comparison because the longer length of the follow-up period allows more time for deterioration of the program effect. On the other hand, the longer follow-up for studies of elementary students tends to equalize the studies in terms of prevalence of substance use at the time of the follow-up. A common criticism of studies of elementary school students using immediate follow-up is that elementary school students are not yet sufficiently at-risk for substance use and therefore the evaluations are not able to detect program effects. Critics have argued that longer follow-ups are necessary to detect the real effects of such programs.

Additional analyses were conducted to explore the possible influence of any confounding of differing follow-up lengths with student grade on effect size. First, results were examined for the effect size from the longest and shortest follow-up time points. This analysis replicates the findings shown in Table 1, indicating that these findings are not a function of effect sizes having been averaged across multiple time points. Second, regression analysis did not reveal a relationship between the length of follow-up period and effect size, suggesting that the observed difference by developmental stage is unlikely to be accounted for by differences in length of follow-up period. Examining the relationship between length of follow-up and effect size only for those studies with late elementary school samples, however, showed a statistically nonsignificant preference for those studies with a follow-up period equal to or greater than 2 years (0.02 vs. 0.08, for shorter and longer follow-up periods, respectively). The mean effect size for the studies with longer follow-up periods of late elementary school students approaches the size of the mean effect for all junior/middle high school studies. In other words, it does not appear that providing prevention programming to elementary school students has longer term prevention benefit that exceeds the benefits of providing such programming to middle school students.

It is worth noting that the confidence interval for the mean effect size for studies with samples of senior high students extends below zero. For this collection of 14 studies, the evidence is insufficient to conclude that school-based prevention programs are effective during this developmental stage. The average effect is comparable to that of the late elementary students,

Table 1. Mean AOD Effect Size for Elementary, Middle/Junior, and Senior High School Aged Students

School grade	Mean ES	95% confidence interval		k^a
		Lower	Upper	
Late elementary	0.05	0.00	0.10	40
Middle/junior	0.09	0.05	0.14	67
Senior	0.04	-0.05	0.14	14

Note. 15 of the 136 studies could not be classified exclusively into one of these three categories and are therefore not included in this analysis. $Q_{\text{BETWEEN}} = 2.16$, $df = 2$, $p > .05$.

^aNumber of effect sizes.

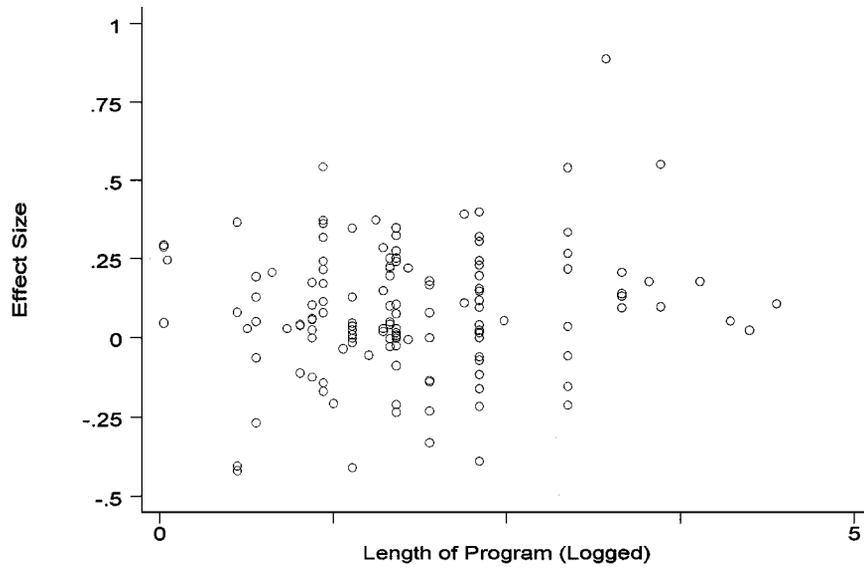


Fig. 1. Scatter plot of length of program logged and effect size.

whose confidence interval also includes zero, and both are quite small in practical terms.

Of course, these results with respect to optimal age for prevention programming do not take into consideration the type of prevention programming. Although it is not possible to examine with existing studies, elementary or high school programs of a different type than have typically been studied might yield higher effect sizes in these age groups.

Characteristics of the Program

Our previous work in this area (Gottfredson *et al.*, 2002; Wilson *et al.*, 2001) summarized earlier focused on the effects across different treatment modalities, such as instructionally based self-control or social competency programs using cognitive-behavioral methods and the general focus of the programs (environmental vs. individual change approaches). Below we focus on the relationship between other aspects of a program's characteristics and program effects: program duration and who delivers the intervention.

The length of the evaluated interventions varied greatly, ranging from less than 1 month to 84 months, not counting summers. Although the modal program length was 9 months, or a full school year, half of the programs were implemented over a period of less than $4\frac{1}{2}$ months. A random effects inverse variance weighted regression analysis found a small positive,

statistically nonsignificant relationship between the length of the intervention (measured in months) and the observed effect size (unstandardized $B = 0.03$, $p = .06$, using logged length of program). The scatter plot in Fig. 1 shows the relationship between length of the program (on a logged scale) and effect size. As can be seen, the positive relationship is driven in large part by a single outlier.⁶ The relationship remains nonsignificant without this outlier. Thus, the evidence suggests that program length is not predictive of student outcomes.

Duration of prevention services is also related to the type of prevention service provided. As in other similar analyses, the number of studies on which modality-specific analyses are based is small, but examining this relationship separately for the five modalities with 10 or more effect sizes we also found no relationship between program duration and effect size with the exception of a miscellaneous category of "other instructional programs." The diversity of programs in this residual category prevents any meaningful conclusion regarding duration of instructional

⁶The outlier evident in this figure was for a 2-year intervention with a booster involving careful behavioral monitoring and reinforcement (Bry, 1982; Bry & George, 1979, 1980). This study found effect size of 0.89 on a measure of substance use. Results from this rigorous study are credible, but the intervention was unusually intensive, involving weekly contact with program staff over a 2-year period as well as frequent communication with the students' parents.

Table 2. Frequency and Percentage of Programs Delivered by Various Types of Individuals

Individual	Involved		Not involved	
	Mean ES	k^a	Mean ES	k^a
Teacher	0.07	75	0.08	61
Peer	0.06	25	0.07	111
Police	0.05	21	0.08	115
Researcher	0.02	15	0.08	121
Mental health/health professional	0.04	8	0.07	128
College student	0.03	2	0.08	131

Note. Programs may be delivered by person in one or more of these categories. None of these differences were statistically significant. ^aNumber of effect sizes.

programs and outcome. We find no compelling evidence that longer programs are more effective than shorter programs.

Not surprisingly, the majority of school-based programs are delivered by teachers (see Table 2). Peers and police officers are the next most common group of individuals who deliver programs. Of interest to this synthesis is whether effectiveness varies as a function of who delivers the program, as suggested by Tobler and Stratton (1997). Table 2 shows that there are no differences in the effect sizes for programs delivered with and without teachers, peers, and police. Although there are sizeable differences for programs delivered by researchers and college students relative to programs delivered by others, these differences are based on a small number of studies and are not statistically significant. But these figures may mask important differences because in many cases, programs are delivered jointly by more than one of the categories of personnel included in Table 2, and also because delivery personnel vary by program modality.

There was a sufficient number of effect sizes for programs delivered by teachers, peers, and police to perform more detailed analyses. A multivariate regression analysis regressed effect size on the degree of involvement of teachers, police, and peers (each a 3-point scale) and an interaction term computed by multiplying the teacher and peer involvement variables (interactions with police delivery were not investigated because police most often deliver programs solo). Also included in the model was a set of dummy codes for the major program modalities. The analysis revealed a significant positive main effect of peer involvement ($p = .05$), no main effect of teacher or police involvement ($p = .71$ and $.64$, respectively), and an interaction between peer and teacher involvement ($p = .02$).

Table 3. Mean AOD Effect Size for Peers and Teacher Involvement in Program Delivery

Peer/teacher involved?	Mean ES	95% confidence interval		k^a
		Lower	Upper	
No peer involvement	0.05	0.02	0.08	71
Peers with teachers	0.04	-0.03	0.11	10
Peers without teachers	0.20	0.09	0.30	8

Note. Analysis included 89 effect sizes. Included in this analysis were those program modalities that had studies in each of the above categories. $Q_{\text{BETWEEN}} = 7.06$, $df = 2$, $p = .03$.

^aNumber of effect sizes.

To understand the nature of the teacher by peer interaction, we limited the analysis to only those program categories that contained studies that involved peers in delivery. Eighty-nine treatment-comparison contrasts from the instructional and normative change program categories were examined. These studies were categorized according to the type of teacher/peer delivery as follows: peers delivered program alone; peers delivered program with teachers; peers not involved in program delivery. Findings from these analyses are presented in Table 3. Effect sizes differ significantly across these categories ($p = .03$). Peer-alone delivery has the highest effect size (0.20), and this effect size is significantly different from zero. The effect size was small and statistically nonsignificant for programs coled by teachers and peers (0.04). Programs run without peer involvement had a similarly small effect size (0.05), but with 71 cases this effect is statistically significantly different from zero ($p < .05$). Thus, these findings suggest a positive benefit of the involvement of peers as leaders, but this benefit disappears when teachers are also involved. Of course, these differences may be confounded with unmeasured features of these programs other than who provided treatment.

Characteristics of the Methodology

The robustness of the above findings may be compromised owing to the variability in methodological characteristics across studies or owing to the selection of studies included in this synthesis (e.g., publication selection bias; Begg, 1994). Any analysis of differences in the pattern of effect sizes across studies is inherently correlational and study characteristics are often confounded, making interpretation of observed differences difficult. A common source of confounding is methodological differences. If such differences

Table 4. Random Effects Regression Analysis for the Effect Sizes Regressed on Method Characteristics ($k^a = 136$)

Variable	Unstandardized <i>B</i>
Random assignment	-0.05
Observed pretest differences	0.07
Students assigned to conditions	0.05
Students unit of analysis	-0.01
Careful selection of measures	0.03
Overall method rating	-0.03
Initial group similarity	0.02
Sample attrition	0.07
Intercept	0.06
R^2	0.04

Note. None of the regression coefficients are statistically significant at $p \leq .05$. $Q_{\text{MODEL}} = 7.02$, $df = 8$, $p = 0.53$.

^aNumber of effect sizes.

are related to the observed effects and confounded with the substantive study characteristics of interest, such as gender of the student sample, then differences due to method may be misinterpreted as differences due to substance.

To address this potential threat, a sensitivity analysis of the above findings was conducted by first regressing the observed effect size on a set of variables that captured important methodological differences across the studies (see Table 4). These included whether the study used random assignment to conditions, observed pretest differences between the comparison and treatment conditions, assigned students to conditions rather than classrooms or schools, used students as the unit-of-analysis, attended to the validity and reliability of the dependent measure(s), were rated highly on a 5-point qualitative index of methodological quality, were rated highly on a 7-point qualitative index of the similarity between the treatment and comparison groups, and the degree of sample attrition. As shown in the table, none of these variables had a statistically significant relationship with the study effect size. Overall, the regression model was not significant and explained very little variability in the effect sizes. A residualized effect size for each study was constructed from this regression analysis (i.e., a covariate adjusted effect size) and the above analyses were rerun. Not surprisingly, all of the above analyses were essentially the same when based on the residualized effect sizes as they had been with the observed effect sizes. Thus, the findings appear robust to these measured methodological differences across studies.

A second issue regarding the robustness of the findings is publication selection bias. This synthesis included unpublished manuscripts as a protection

against this source of bias. Counter to the typical finding in meta-analysis (Lipsey & Wilson, 1993), this study found that studies published in peer-reviewed journals ($n = 91$) and as book chapters ($n = 6$) had a lower mean effect size than did unpublished studies ($n = 39$, mean effect sizes = 0.05, 0.08, and 0.12, respectively). These differences were not statistically significant. Furthermore, a funnel plot (Light *et al.*, 1994) of the relationship between effect size and sample size had the desired funnel shape, suggesting that this collection of studies is unlikely to be affected by publication bias.

CONCLUSIONS

This meta-analysis attempted to identify potential moderator variables that would differentiate more from less effective school-based substance abuse prevention programs. It examined characteristics of the populations served (high risk vs. general population and age of the population), characteristics of the program (total duration and the role of the person delivering the service), and characteristics of the methodology.

Only 11 out of 136 treatment-control contrasts in this study involved target populations that were at elevated risk for substance use or other problem behaviors. This contrasts with an earlier meta-analysis of school-based delinquency prevention programs (Wilson *et al.*, 2001) that found a substantial number of studies involving high-risk youths, and that such programs had larger effect sizes than universal programs. This study found that although in general programs targeting high-risk populations are about as effective as those provided universally, cognitive-behaviorally based prevention programs are more effective at reducing substance use when delivered to high-risk than general populations. But this finding was based on only a handful of studies. Additional studies are needed to explore the relative strengths of selective versus universal prevention programming and to determine whether the additional costs (both in terms of personnel and materials costs and opportunity costs incurred by removing youths from some perhaps more beneficial educational activity) related to universal programming are justified. If our finding holds up to additional scrutiny, it will be consistent with conclusions from studies delinquency prevention programs that targeting high risk populations is more effective than targeting lower risk populations (e.g., Andrews *et al.*, 1990; Wilson & Lipsey, 2000). It will

also support theoretical work in criminology (Moffitt, 1993) that has demonstrated that the offending population consists of two distinct groups: a large group of individuals who experiment with illegal activities for a relatively short period of time during adolescence and then naturally desist, and a smaller group of offenders who begin their offending behavior earlier, continue for a longer period, and offend at higher rates throughout their period of offending. This latter group is responsible for the majority of crime that occurs. If this distribution also describes adolescent substance users, targeting more prevention services towards higher risk youths would make more sense than the current practice of directing most resources to the general population.

The study also found that programs targeting middle school aged youths are slightly more effective than those targeting elementary school aged or senior high school aged youths, but the difference is not statistically significant. The average effect size for programs targeting senior high students was not significantly different from zero, and the confidence interval for programs targeting elementary students also included zero. Only for middle school programs does the evidence clearly imply effectiveness for reducing AOD. The evidence available does not completely rule out the possibility of undetected delay effects for the late elementary school programs, but the fact that the results were unchanged when only the latest available time point was examined argues against such “sleeper” effects. Additional research is necessary, however, to determine the relative benefits of programs that begin in elementary school and continue through the middle grades compared with programs that target one or the other developmental age group.

The analysis of program duration suggested that longer is not necessarily better. Programs of relatively brief duration (less than 4.5 months) are generally as effective as those of longer duration (more than 4.5 months). This finding makes sense in light of the generally small effect sizes found for school-based prevention activities overall. Many programs are not effective or have very small effects. Extending their duration is not likely to produce much in the way of benefits. It is possible that extending the duration of especially effective types of prevention would be beneficial, but analyses done separately by prevention modality did not support this hypothesis. It may also be the case that program length is a poor proxy for program intensity. A more sensitive measure of program intensity may have produced different results.

Finally, the role of the person delivering the preventive intervention is generally not related to the effectiveness of the program, but controlling for type of prevention activity yields a positive effect for peer involvement in delivery. A statistical interaction between teacher and peer involvement in program delivery was also observed, however. Programs that involved *peers alone* in program delivery were most effective. Those that involved no peers or peers in conjunction with teachers were far less effective. This apparent interaction may help to explain seemingly contradictory results in the prevention literature. Although some studies (e.g., Perry *et al.*, 1989) have found that substance abuse prevention programs are more effective when led by peers than by teachers, other studies (e.g., Ellickson & Bell, 1990) found no such advantage for peer-led programs. Also, Tobler’s meta-analysis (Tobler, 1992) found no evidence that programs with peer leaders produce better outcomes than programs of similar content led by adults. Our analysis suggests that the benefits that accrue to peer delivery of a program may disappear when the teacher shares the delivery role with the peer.

This analysis raised several questions that will require additional original research to resolve: The finding that, at least for programs involving social competency instruction, targeting higher risk youths results in higher effect sizes than targeting the general school population raises questions about the utility of continuing the current practice of applying most prevention programming universally. This issue requires additional research, because evaluations of targeted substance abuse prevention programs are scarce. We need first to better understand the features of effective targeted prevention program. It is not safe to assume that programs that have been demonstrated effective with general populations will also work with youths who have already initiated drug use. In fact, evidence suggests that universal programs have no effect (Hansen *et al.*, 1988) or even negative effects (Bell *et al.*, 1993; Ellickson *et al.*, 1993; Ellickson & Bell, 1990) for drug-using populations. Brown and Kreft (1998) have argued that the “no use” messages typically conveyed in universal programs may actually increase use among those most at risk for using because these youths are more knowledgeable about drugs and their effects than the curricula assume. The messages conveyed in the curricula may appear naive to such youths, and may be rejected.

Once the elements of effective prevention programming for high-risk youths are better understood, models that incorporate these features must

be developed and tested and cost/benefit analyses of targeted programs and universal programs conducted to increase understanding of the pay-off in terms of substance abuse prevention relative to the costs for each type of program.

Additional research is also needed to understand the benefits that may accrue as a result of early prevention programming. Our analysis shows that programs targeting middle school students only have larger effects than do those targeting elementary or senior high school students only. This finding is consistent with a recent decision by DARE America to shift its programming upward to the middle grades. However, we do not know what, if anything, can be gained through programs that begin in the elementary school grades and continue through the middle school grades. No study has examined the cumulative effect of elementary plus middle school lessons relative to the effects of either alone. Such research is needed to better understand the most effective timing for prevention programming.

ACKNOWLEDGMENTS

This work was supported in part by a grant from Jerry Lee. We gratefully acknowledge the research assistance of Todd Armstrong, Jennifer Castro, Elaine Eggleston, Paul Hirschfield, Veronica Puryear, John Ridgely, Stacy Skroban Najaka, and Shannon Womer.

REFERENCES

- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., Cullen, F. T., Lab, S. P., & Whitehead, J. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*, 369–404.
- Battistich, V., Schaps, E., Watson, M., & Solomon, D. (1996). Prevention effects of the Child Development Project: Early findings from an ongoing multisite demonstration trial. *Journal of Adolescent Research*, *11*(1), 12–35.
- Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–410). New York: Russell Sage Foundation.
- Bell, R. M., Ellickson, P. L., & Harrison, E. R. (1993). Do drug prevention effects persist into high school? How Project ALERT did with ninth graders. *Preventive Medicine*, *22*, 463–483.
- Botvin, G. J. (1990). Substance abuse prevention: Theory, practice, and effectiveness. In M. Tonry & J. Q. Wilson (Eds.), *Drugs and crime* (pp. 461–519). Chicago, IL: University of Chicago Press.
- Botvin, G. J., Baker, E., Renick, N. L., Filazzola, A. D., & Botvin, E. M. (1984). A cognitive-behavioral approach to substance abuse prevention. *Addictive Behaviors*, *9*, 137–147.
- Botvin, G. J., Schinke, S., & Orlandi, M. A. (1995). School-based health promotion: Substance abuse and sexual behavior. *Applied and Preventive Psychology*, *4*, 167–184.
- Brown, J. H., & Kreft, I. G. G. (1998). Zero effects of drug prevention programs: Issues and solutions. *Evaluation Review*, *22*(1), 3–14.
- Bry, B. H. (1982). Reducing the incidence of adolescent problems through preventive intervention: One- and five-year follow-up. *American Journal of Community Psychology*, *10*(3), 265–276.
- Bry, B. H., & George, F. E. (1979). Evaluating and improving prevention programs: A strategy from drug abuse. *Evaluation and Program Planning*, *2*, 127–136.
- Bry, B. H., & George, F. E. (1980). The preventive effects of early intervention on the attendance and grades of urban adolescents. *Professional Psychology*, *11*, 252–260.
- Comer, J. P. (1985). The Yale-New Haven Primary Prevention Project: A follow-up study. *Journal of the American Academy of Child Psychiatry*, *24*(2), 54–160.
- Committee on Data and Research for Policy on Illegal Drugs (2001). *Informing America's policy on illegal drugs: What we don't know keeps hurting us*. Washington, DC: National Academy Press.
- Cook, T. D., Hunt, H. D., & Murphy, R. F. (1998). *Comer's school development program in Chicago: A theory-based evaluation*. Chicago, IL: Institute for Policy Research, Northwestern University.
- Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., Laudolf, J., Turkkan, J. S., Ford, C., & Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology*, *14*, 317–345.
- Dryfoos, J. G. (1990). *Adolescents at risk: Prevalence and prevention*. New York: Oxford University Press.
- Durlak, J. A. (1995). *School-based prevention programs for children and adolescents*. Thousand Oaks, CA: Sage.
- Eggert, L. L., Thompson, E. A., Herting, J. R., Nicholas, L. J., & Dicker, B. G. (1994). Preventing adolescent drug abuse and high school dropout through an intensive school-based social network development program. *American Journal of Health Promotion*, *8*(3), 202–215.
- Ellickson, P. L., & Bell, R. M. (1990). Drug prevention in junior high: A multi-site longitudinal test. *Science*, *247*, 1299–1305.
- Ellickson, P. L., Bell, R. M., & McGuigan, K. (1993). Preventing adolescent drug use: Long-term results of a junior high program. *American Journal of Public Health*, *83*(6), 856–861.
- Ennett, S. T., Tobler, N. S., Ringwalt, C. L., & Flewelling, R. L. (1994). How effective is Drug Abuse Resistance Education? A meta-analysis of Project D.A.R.E. outcome evaluations. *American Journal of Public Health*, *84*, 1394–1401.
- Gerstein, D. R., & Green, L. W. (Eds.) (1993). *Preventing drug abuse: What do we know?* Washington, DC: National Academy Press.
- Gorman, D. M. (1995). Are school-based resistance skills training programs effective in preventing alcohol misuse? *Journal of Alcohol and Drug Education*, *41*, 74–98.
- Gottfredson, D. C. (1997). School-based crime prevention. In L. W. Sherman, D. C. Gottfredson, D. MacKenzie, J. Eck, P. Reuter, & S. Bushway (Eds.), *Preventing crime: What works, what doesn't, what's promising: A report to the United States Congress*. Washington, DC: U.S. Department of Justice, Office of Justice Programs.
- Gottfredson, D. C. (2001). *Schools and delinquency*. New York: Cambridge University Press.
- Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency*, *39*(1), 3–35.
- Gottfredson, D. C., Wilson, D. B., & Najaka, S. S. (2002). School-based crime prevention. In D. P. Farrington, L. W. Sherman, B. Welsh, & D. L. Mackenzie (Eds.), *Evidence-based crime prevention*. United Kingdom: Routledge.

- Gottfredson, G. D., & Gottfredson, D. C. (2001). What schools do to prevent delinquency and promote safe environments. *Journal of Educational and Psychological Consultation, 12*(3), 313–344.
- Hansen, W. B., (1992). School-based substance abuse prevention: A review of the state of the art in curriculum: 1980–1990. *Health Education Research, 7*, 403–430.
- Hansen, W. B., & Graham, J. W. (1991). Preventing alcohol, marijuana, and cigarette use among adolescents: Peer pressure resistance training versus establishing conservative norms. *Preventive Medicine, 20*, 414–430.
- Hansen, W. B., Johnson, C. A., Flay, B. R., Graham, J. W., & Sobel, J. (1988). Affective and social influences approaches to the prevention of multiple substance abuse among seventh grade students: Results from Project SMART. *Preventive Medicine, 17*, 135–154.
- Hansen, W. B., & O'Malley, P. M. (1996). Drug use. In R. J. DiClemente, W. B. Hansen, & L. E. Ponton (Eds.), *Handbook of adolescent health risk behavior* (pp. 161–192). New York: Plenum Press.
- Hawkins, J. D., Arthur, M. W., & Catalano, R. F. (1995). Preventing substance abuse. In M. Tonry & D. Farrington (Eds.), *Building a safer society: Strategic approaches to crime prevention* (pp. 343–427). Chicago, IL: University of Chicago Press.
- Huizinga, D., & Jacob-Chien, C. (1998). The contemporaneous co-occurrence of serious and violent juvenile offending and other problem behaviors. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 47–67). Thousand Oaks, CA: Sage.
- Institute of Medicine (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.
- Kellam, S. G., & Anthony, J. C. (1998). Targeting early antecedents to prevent tobacco smoking: Findings from an epidemiologically based randomized field trial. *American Journal of Public Health, 88*, 1490–1495.
- Kellam, S. G., Rebok, G. W., Ialongo, N., & Mayer, L. S. (1994). The course and malleability of aggressive behavior from early first grade into middle school: Results of a developmental epidemiologically-based preventive trial. *Journal of Child Psychology and Psychiatry, 35*, 259–281.
- Light, R. J., Singer, J. D., & Willett, J. B. (1994). The visual presentation and interpretation of meta-analyses. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–410). New York: Russell Sage Foundation.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*, 1181–1209.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage.
- Lochman, J. E. (1992). Cognitive-behavioral intervention with aggressive boys: Three-year follow-up and preventive effects. *Journal of Consulting and Clinical Psychology, 60*, 426–432.
- Moffitt, T. E. (1993). Adolescence-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674–701.
- Norman, E., & Turner, S. (1993). Adolescent substance abuse prevention programs: Theories, models, and research in the encouraging 80's. *Journal of Primary Prevention, 14*, 3–20.
- Perry, C. D., Grant, M., Ernberg, G., Florenzano, R. U., Langdon, M. C., Myeni, A. D., Waahlberg, R., Berg, S., Andersson, K., Fisher, J., Blaze-Temple, D., Cross, D., Saunders, B., Jacobs, D. R., & Schmid, T. (1989). WHO collaborative study on alcohol education and young people: Outcomes of a four-country pilot study. *The International Journal of the Addictions, 24*, 1145–1171.
- Perry, C. L., Williams, C. L., Veblen-Mortenson, S., Toomey, T. L., Komro, K. A., Anstine, P. S., McGovern, P. G., Finnegan, J. R., Forster, J. L., Wagenaar, A. C., & Wolfson, M. (1996). Project Northland: Outcomes of a communitywide alcohol use prevention program during early adolescence. *American Journal of Public Health, 86*, 956–965.
- Robins, L. N., & Przybeck, T. R. (1985). Age of onset of drug use as a factor in drug and other disorders. In C. L. Jones & R. J. Battjes (Eds.), *Etiology of drug abuse: Implications for prevention* (National Institute on Drug Abuse Research Monograph No. 56). Washington, DC: U.S. Government Printing Office.
- Tobler, N. S. (1992). Drug prevention programs can work: Research findings. *Journal of Addictive Diseases, 11*(3), 1–28.
- Tobler, N. S., & Stratton, H. H. (1997). Effectiveness of school-based drug prevention programs: A meta-analysis of the research. *Journal of Primary Prevention, 18*(1), 71–128.
- Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2001). School-based prevention of problem behaviors: A meta-analysis. *Journal of Quantitative Criminology, 17*(3), 247–272.
- Wilson, S. J., & Lipsey, M. W. (2000). *Effects of school violence prevention programs on aggressive and disruptive behavior: A meta-analysis of outcome evaluations*. Unpublished manuscript, Vanderbilt University, Nashville, TN.