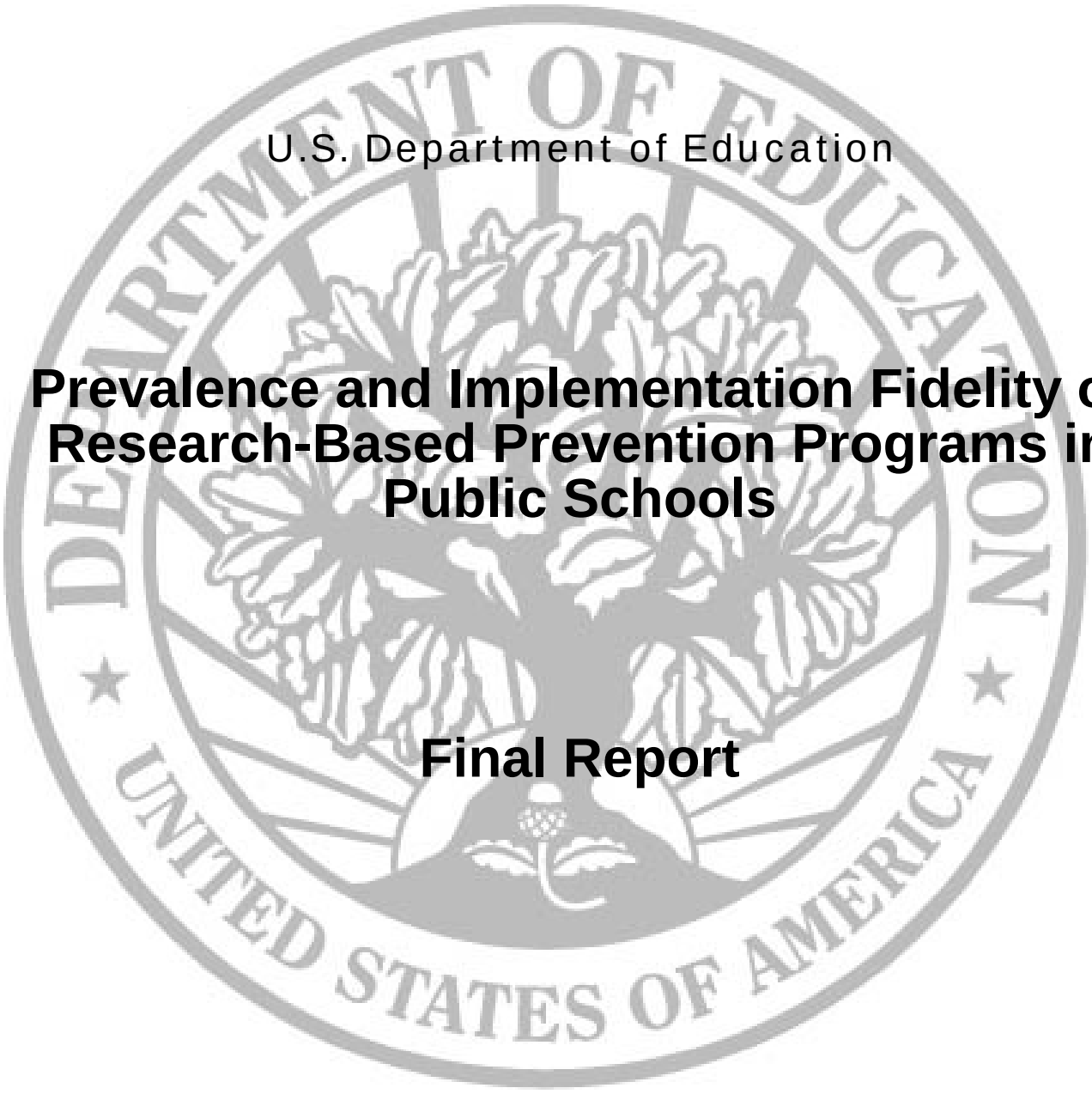

U.S. Department of Education

Prevalence and Implementation Fidelity of Research-Based Prevention Programs in Public Schools

Final Report

Prevalence and Implementation Fidelity of Research-Based Prevention Programs in Public Schools

Final Report

Scott Crosse
Barbara Williams
Carol A. Hagen
Michele Harmon
Liam Ristow
Ralph DiGaetano
Pamela Broene
Debbie Alexander
Margaret Tseng
James H. Derzon

Prepared for:
U.S. Department of Education
Office of Planning, Evaluation and Policy Development
Policy and Program Studies Service

Prepared by:
Westat
Rockville, Maryland

2011

This report was prepared for the U.S. Department of Education under Contract No. ED-00-CO-0119. The project monitor was Michael Fong in the Policy and Program Studies Service. The views expressed herein are those of the contractor. No official endorsement by the U.S. Department of Education is intended or should be inferred.

U.S. Department of Education
Arne Duncan
Secretary

Office of Planning, Evaluation and Policy Development
Carmel Martin
Assistant Secretary

Policy and Program Studies Service
Stuart Kerachsky
Director

December 2011

This report is in the public domain. Authorization to produce it in whole or in part is granted. Although permission to reprint this publication is not necessary, the citation should be: U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service, *Prevalence and Implementation Fidelity of Research-Based Prevention Programs in Public Schools: Final Report*, Washington, D.C., 2011.

This report is available on the Department's Web site at
http://www.ed.gov/about/offices/list/opepd/ppss/reports.html.

On request, this publication is available in alternative formats, such as Braille, large print or CD. For more information, please contact the Department's Alternate Format Center at 202-260-0852 or 202-260-0818.

Content Contact
Michael Fong
202-401-7462
michael.fong@ed.gov

Contents

Chapter		Page
Executive Summary		xv
1	Introduction	1
	Prevalence of Youth Alcohol, Tobacco, and Other Drug Use and School Crime	1
	Youth ATOD use	1
	School crime	2
	SDFSCA Program and Efforts to Address Youth ATOD Use and School Crime	3
	Previous Research	4
	Prevalence of research-based programs	4
	Implementation fidelity of research-based programs	6
	Purpose of the Study	8
	Organization of Report	9
2	Methodology	11
	Sample Design	11
	Measurement	13
	Identification of research-based programs and practices	13
	Data collection instruments	16
	Development of implementation fidelity measures	17
	Data Collection and Processing	19
	Analysis	20
	Limitations	21
3	Prevalence of Prevention Programs	25
	Prevention Programs Overall	25
	Number of programs	25
	Types of programs	26
	Frequently Reported Programs	27
	Research-Based Programs	29
	Number of programs	29
	Types of programs	30
	Frequently reported programs	30

Contents (continued)

Chapter		Page
4	Implementation Fidelity Of Research–Based Prevention Programs	31
	Research-Based Programs Included in Fidelity Analyses	31
	Program-Specific Measures and Generic Measures	33
	Achievement of Program-Specific Standards	35
	Achievement of Generic Standards	37
	Overall Implementation Fidelity	39
5	Factors Associated with Implementation Fidelity of Research-Based Prevention Programs	41
	School Characteristics	42
	School urbanicity	44
	School enrollment	45
	School percent eligible for free or reduced-price lunches	46
	Training on Programs	47
	Quality of initial training provided	47
	Amount of training provided	50
	District Characteristics	52
	Federal and State Support and Policies	53
	State provision of training or technical assistance on implementing programs	53
	State provision of training or technical assistance on evaluating programs	54
6	Conclusions	57
	Prevalence of Research-Based Programming	57
	Implementation Fidelity of Research-Based Programming	58
	Limitations Pertaining to Potential Misreporting	60
	References	61

Contents (continued)

Appendixes

Appendix		Page
A	Methodology	65
B	Summary Report on the research synthesis	123
C	Efforts to Collect Data on Program Funding From School Personnel	143

Contents (continued)

Tables

Table		Page
1	Schools by number of prevention programs and instructional level: 2004–05	26
2	Prevention programs by program type and instructional level: 2004–05	27
3	Most frequently reported named prevention programs: 2004–05	29
4	Most frequently reported research-based programs: 2004–05	30
5	Research-based curriculum programs in fidelity analyses by instructional level: 2004–05	32
6	Program-specific and generic measures and standards by program	34
7	Research-based curriculum programs passing on program-specific standards by instructional level: 2004–05	36
8	Research-based curriculum programs passing on program-specific standards by program: 2004–05	37
9	Research-based curriculum programs passing on generic standards by instructional level: 2004–05	38
10	Research-based curriculum programs passing on generic standards by program: 2004–05	39
11	Overall implementation fidelity of research-based curriculum programs by instructional level: 2004–05	40
12	Overall implementation fidelity of research-based curriculum programs by program: 2004–05	40
13	Summary of statistically significant relationships between implementation fidelity standards, and program, school, and district characteristics: 2004–05	43
14	Research-based curriculum programs passing on individual program-specific standards by school urbanicity: 2004–05	44

Contents (continued)

Tables (continued)

Table		Page
15	Overall implementation fidelity of research-based curriculum programs by school urbanicity: 2004–05	45
16	Research-based curriculum programs passing on individual program-specific standards by school enrollment: 2004–05	46
17	Overall implementation fidelity of research-based curriculum programs by school enrollment: 2004–05	46
18	Research-based curriculum programs passing on individual generic standards by school percent eligible for free or reduced-price lunches: 2004–05	47
19	Research-based curriculum programs passing on individual program-specific standards by quality of initial training: 2004–05	48
20	Research-based curriculum programs passing on individual generic standards by quality of initial training: 2004–05	49
21	Overall implementation fidelity of research-based curriculum programs by quality of initial training: 2004–05	50
22	Research-based curriculum programs passing on individual program-specific standards by amount of training: 2004–05	51
23	Research-based curriculum programs passing on individual generic standards by amount of training: 2004–05	52
24	Overall implementation fidelity of research-based curriculum programs by amount of training: 2004–05	52
25	Research-based curriculum programs passing on individual generic standards, by district percent eligible for free or reduced-price lunches: 2004–05	53
26	Research-based curriculum programs passing on individual program-specific standards by whether state-provided training or technical assistance on implementing programs: 2004–05	54

Contents (continued)

Tables (continued)

Table		Page
27	Research-based curriculum programs passing on individual program-specific standards by whether state-provided training or technical assistance on evaluating programs: 2004–05	55
A-1	Grade-span definitions of school classification for sampling purposes	69
A-2	Sample sizes of schools across the 27 strata	71
A-3	Variation in sample sizes of programs per school	72
A-4	Response for district recruitment, by geographical region and metropolitan status	82
A-5	Response for district survey, by geographical region and metropolitan status	83
A-6	Response for prevention program survey, by geographical region and metropolitan status	84
A-7	Response for provider survey, by geographical region and metropolitan status	85
A-8	Response for principal survey, by geographical region and metropolitan status	86
A-9	Frequency distribution across the general response disposition categories for the Provider Questionnaire	98
A-10	Standards for measuring fidelity of implementation	111
C-1	Prevention programs by whether they received funding from SDFSCA Program and instructional level: 2004–05	145
C-2	Prevention programs by whether research-based and received funding from SDFSCA program: 2004–05	145
C-3	Research-based curriculum programs in fidelity analyses by whether received funding from SDFSCA Program: 2004–05	147

Contents (continued)

Tables (continued)

Table		Page
C-4	Research-based curriculum programs passing on program-specific standards by whether they received funding from SDFSCA Program: 2004–05	148
C-5	Research-based curriculum programs passing on generic standards by whether they received funding from SDFSCA Program: 2004–05	149
C-6	Overall implementation fidelity of research-based curriculum programs by whether they received funding from SDFSCA Program: 2004–05	149

Contents (continued)

Figure

Figure		Page
B-1	Summary of research review	125

Contents (continued)

Exhibits

Exhibit		Page
1	Study questions	8
B-1	Lists of promising and effective programs	126
B-2	Prevention programs aligned with the mission of the Safe and Drug-Free Schools and Communities Program	127
B-3	Status of programs that remained after initial screening for relevance to study	128
B-4	Expert panel members	135
B-5	Descriptions of effective programs resulting from research synthesis	137

EXECUTIVE SUMMARY

This report presents descriptive information about the prevalence and quality of implementation of research-based programs from the Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime. The study found that, while schools reported implementing a large number of prevention programs during the 2004–05 school year, only a small proportion reported implementing programs (approximately 7.8 percent) supported by research evidence that met the methodological standards established by the study. It also estimated that less than one-half of those implementing research-based curriculum prevention programs (approximately 44.3 percent of the estimated 7.8 percent of research-based programs) met minimal standards for fidelity of implementation during the 2004–05 school year.[1] Given the proportion of prevention programs offered in schools that are research-based, the estimated proportion of all curriculum prevention programs that are research-based and well-implemented is approximately 3.5 percent. A variety of factors are associated with the adoption and fidelity of implementation of research-based prevention programs.

Main Findings

- Information was examined on over 300 programs that were found on existing lists of "promising" or "effective" prevention programs. The study identified 19 school-based prevention programs that demonstrated evidence of effectiveness through this systematic review of literature.

- A survey collecting data on prevention programs in the nation's public schools found that the 19 research-based programs accounted for an estimated 7.8 percent of the programs implemented during the 2004–05 school year.

- Approximately 44.3 percent of the research-based curriculum programs, or just 3.5 percent of all programs implemented in schools, met minimum standards for overall fidelity of implementation based on four program-specific measures.

[1] Curriculum programs involve the provision of training or instruction to students. Results from Phase 1 indicate that they account for approximately 97 percent of the research-based prevention programs delivered in schools during the 2004–05 school year.

Study Background

School-based prevention programs receive support from a variety of sources. Before Fiscal Year (FY) 2010, the Safe and Drug-Free Schools and Communities Act (SDFSCA) Program distributed formula grants through a State Education Agency (SEA) Program and Governors' Program.[2] As a part of the national effort to provide programming in elementary and secondary schools, the SDFSCA Program provided funding to states to support substance abuse and violence prevention programs.

The Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime was funded by the U.S. Department of Education to measure the prevalence of research-based programs in schools intended to prevent youth substance abuse and school crime and to assess the fidelity of implementation of those research-based programs. These were program performance measures under the SEA Program. Phase 1 of the study (Prevalence Study) focused on meeting the first objective; and Phase 2 (Fidelity Study) focused on meeting the second objective. The study defined a prevention program as a school intervention or initiative that aims to mitigate or eliminate negative student behaviors like drug use and school violence.

Because not every prevention program is supported with research, one major component of the overall study was to identify research-based prevention programs based on a systematic review of the research literature. Developing a valid and useful list of programs for the study entailed compiling and screening existing lists of research-based prevention programs, assessing the quality of evidence on the programs that pass the screens, and making judgments on whether high-quality evidence on a given program indicates a pattern of program-related effects. To identify specific programs, we identified and reviewed over 2,000 individual study reports on programs that were judged to be effective by external sources (a total of 317 prevention programs entered this review process). The study identified 19 school-based prevention programs that demonstrated evidence of effectiveness through this systematic review of literature. The evidence of effectiveness must have included quantitative and statistically significant results on at least one behavioral outcome from an experimental or strong quasi-experimental design.

[2] SDFSCA Program refers to the formula grant program that was administered by the U.S. Department of Education and implemented in school districts and communities through SEAs and governors' offices. Funding for the SDFSCA Program was eliminated beginning in FY 2010.

During spring and fall 2006, nationally representative surveys of 2,500 public school districts and 5,847 public schools were conducted for the Study. School principals reported on the prevention programs operating in their schools during the 2004–05 school year, including the research-based programs that the Prevalence Study identified through an extensive critical review of the research literature. The survey response rate was 86 percent. District coordinators provided information on aspects of their prevention programming, including those that were potentially associated with the adoption by schools of research-based programs. The survey response rate was 91 percent. Program coordinators in the schools with research-based programs reported on their implementation of those programs and of research-based practices. The survey response rate was 78 percent. Reports on aspects of implementation of curriculum programs were compared to implementation standards, some of which draw upon program developers' specifications for implementation.

Although this study has many strengths, it also has limitations. The research review used to develop the list of research-based programs for the study may have inadvertently excluded some research from consideration. Additional limitations pertain to data quality, which was affected by recall problems experienced by respondents, overreporting of some programs, and missing data on SDFSCA Program funding for prevention programs. Finally, the Fidelity Study had limitations, pertaining to the application of program-specific standards for assessing the fidelity of program implementation, that could lead to underestimates of the fidelity of implementation of the research-based curriculum programs.

In the remainder of this summary, main findings from the Prevalence Study and Fidelity Study are highlighted.

Prevalence of Research-based Prevention Programs

During the 2004–05-school year, schools reported implementing a large number and diverse types of prevention programs. However, only a small proportion of these programs were research-based. This finding also applies to the prevention programs that received funding from the SDFSCA Program.

Prevention Programs Overall

- On average, schools used 9.0 different prevention programs. The number of programs per school used by middle schools (10.1 programs) was somewhat higher than that for elementary and high schools (8.8 and 8.9 programs, respectively). While 14.8 percent of schools reported using no prevention programs, 11.1 percent identified more than 20 programs.

- Curriculum, instruction, and training programs were the most frequently used program type, accounting for 23.6 percent of all programs implemented. The program types least frequently used by schools included improvements to instructional practices (3.6 percent of programs) and youth roles in regulating student conduct (3.7 percent of programs). The types of programs used varied only slightly among instructional levels.

Research-based Prevention Programs Overall

To identify research-based programs, the Prevalence Study completed an extensive critical review of the research literature on school-based programs intended to prevent or reduce youth substance abuse and school crime. Programs were classified as "research-based" if they were supported by studies that met strict methodological standards.[3]

- Only 7.8 percent of the prevention programs in operation during the 2004–05 were research-based. In other words, the average number of research-based programs implemented per school was less than 1 (0.7), while the average number of all prevention programs was 9.0.

- In terms of schools, 40.7 percent of schools implemented at least one research-based program.

- The vast majority of research-based programs were curriculum, instruction, and training programs, accounting for 97.1 percent of all research-based programs implemented.

[3] For the Prevalence Study, 19 programs were classified as research-based. Two additional programs were included in the analyses for the study, in an attempt to reconcile the Prevalence Study list of research-based programs with other lists that were developed subsequently.

Fidelity of Implementation of Research-based Curriculum Prevention Programs

During the 2004–05 school year, less than half (44.3 percent) of the research-based curriculum programs provided in schools met minimum standards for overall fidelity of implementation.[4] The programs performed substantially better on some standards of implementation fidelity than others. To develop the standards on which the program-specific measures are based, researchers obtained and coded aspects of the program implementation manuals and other materials for each of the research-based programs. These measures were selected because they reflect important aspects of implementation that have been discussed in the research literature, and they were supported by high-quality data (i.e., low item nonresponse, adequate distribution of responses, and clear basis for developing standards).

Curriculum Prevention Programs Overall

The Fidelity Study examined the quality of program implementation in terms of four program-specific fidelity measures and two generic fidelity measures.[5]

- Approximately 44.3 percent of programs passed on all four program-specific fidelity measures, which was the minimum threshold set for adequate overall fidelity of program implementation.[6]
- A large proportion of programs met fidelity standards on topics covered (92.0 percent) and level of risk targeted (88.8 percent).
- On measures related to the amount of classroom time allocated to prevention programming, number of lessons and frequency of student participation, slightly less than two-thirds of programs passed (63.4 percent and 64.8 percent, respectively).

---

[4] For the Fidelity Study, data on aspects of implementation were collected on 19 of the 21 programs included in the Prevalence Study. A sufficient number of responses were obtained on 10 of those 19 programs to warrant including them in the Fidelity Study analyses. Those 10 programs are all curriculum programs.

[5] The study identified six measures of program fidelity, four are "program specific" as identified by the program's instruction manuals or literature and two measures are "generic" based on best practices that can be applied to all programs. The four program-specific fidelity measures are: (1) topics covered, (2) level of risk targeted, (3) number of lessons, and (4) frequency of student participation. The two generic fidelity measures are (5) instructional strategies, and (6) rewards, recognition, and mastery assessment.

[6] The study's researchers established thresholds or cut points for deciding whether the implementation of a given program met the standard. For the purposes of this study, achievement of all four program-specific standards was the minimum threshold for adequate overall implementation fidelity. These program-specific measures represent key design features of the program, and therefore each program-specific measure must be met.

- Programs performed relatively well on instructional strategies (77.3 percent passed) and poorly on rewards, recognition, and mastery assessment (32.6 percent passed).

Factors Associated with Fidelity of Implementation of Research-based Curriculum Prevention Programs

A number of factors were associated with whether programs met fidelity standards. Some of these factors may have implications for practice.

School and District Characteristics

The following characteristics may indicate the types of schools and districts to target in efforts to increase fidelity of implementation of research-based curriculum programs:

- Urban and suburban schools (41.9 percent of programs in urban schools and 38.1 percent of programs in suburban schools were implemented with adequate overall fidelity vs. 51.3 percent of programs in rural schools); and

- Schools with larger enrollments (43.7 percent of programs in schools with 300 to 999 students and 29.7 percent of programs in schools 1,000 or more students were implemented with adequate overall fidelity vs. 53.7 percent of programs in schools with less than 300 students).

District prevention programming factors

Relatively few factors pertaining to prevention programming were associated with fidelity of implementation. These programming factors point to the importance of training on curriculum programs for achieving high quality implementation. They include the following:

- Quality of initial training on curriculum programs (61.4 percent of programs with high quality training were implemented with adequate overall quality vs. 43.7 percent and 33.3 percent of programs with medium or low quality training, respectively). See Appendix A for details on measuring training quality through a composite ratings scale; and

- Amount of training, including ongoing training (62.5 percent of programs with a high amount of training and 50.0 percent of those with a medium amount were implemented with adequate overall quality vs. 36.1 percent of programs with a low amount of training).

Conclusions

Much greater emphasis is needed on the adoption of research-based prevention programs in schools and on the fidelity of implementation of those programs. This conclusion is based on the low proportion of prevention programs operating in schools that are research-based (approximately 7.8 percent of all prevention programs) and on the low proportion of research-based curriculum programs operating in schools that meet fidelity standards (approximately 44.3 percent of all research-based curriculum programs). A limited set of findings on school and district characteristics associated with the fidelity of implementation of research-based programs may point to how to target efforts to boost fidelity; findings on programming factors may indicate the types of approaches that could be effective in increasing the fidelity of implementation of research-based programs.

To improve the prevalence of research-based prevention programs, states and districts should be accountable for meeting a set of program requirements that include using program funds only for prevention programs that demonstrate evidence of effectiveness. In addition, implementation standards should be identified and emphasized to help states and districts develop better programming strategies and provide additional support and training to enhance the overall quality of implementation. Doing less could mean that the vast majority of school-based programs will fail to meet the goals of preventing youth substance abuse and school crime.

1. INTRODUCTION

This report presents findings from the first and second phases of the Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, which was sponsored by the U.S. Department of Education. The first phase focused on the extent to which research-based prevention programs (intended to prevent or reduce youth alcohol, tobacco, and other drug use, and school crime) are operating in public elementary and secondary schools. The second phase examined the program implementation fidelity of those research-based programs. This chapter discusses (a) the prevalence of youth alcohol, tobacco, and other drug use, as well as school crime; (b) the role of the Safe and Drug-Free Schools and Communities Act (SDFSCA) Program in efforts to lower the levels of such problem behavior; (c) the previous research on the adoption of and fidelity of implementation of prevention programs; (d) the purpose of the Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime (here after referred to as "the study"); and (e) the organization of the report.

Prevalence of Youth Alcohol, Tobacco, and Other Drug Use and School Crime

The goal of many school-based prevention programs is to prevent or reduce youth alcohol, tobacco, and other drug (ATOD) use, and school crime. Among the many concerns that are associated with these problems, of special importance to educators is the extent to which youth ATOD use and school crime can hamper efforts to create school environments that are conducive to learning. While levels of youth ATOD use and school crime have been declining in recent years, they remain unacceptably high.

Youth ATOD use. Based on recent results from the Monitoring the Future Study, use of alcohol, tobacco, and other drugs by high school students has continued a pattern of decline that began within the last decade (Johnson, O'Malley, Bachman, and Schulenberg, 2009). In 2008, approximately 22.3 percent of 12th-grade students used any illicit drug in the last 30 days; it was 15.8 percent and 7.6 percent for 10th-grade students and 8th-grade students, respectively. These rates are down by approximately 4 to 6 percentage points from the levels reported in 1999.

A similar pattern occurs for any alcohol use in the last 30 days. In 2008, the figures were 43.1 percent for 12th-grade students compared to 51.0 percent in 1999; 28.8 percent for 10th-grade students, down from 40.0 percent in 1999; and 15.9 percent for 8th-grade students, down from 24.0 percent in 1997. In 2008, the estimated percentage of students drinking to get drunk in the last 30 days was 27.6 for 12th-grade students, 14.4 for 10th-grade students, and 5.4 for 8th-grade students; again, these levels are lower than they were in 1999.

Although the decreases in use over the past decade are perhaps more dramatic for cigarette use (as much as 14 percentage points), the levels of use in the past 30 days are still high. In 2008, it was estimated to be 20.4 percent for 12th-grade students, 12.3 percent for 10th-grade students, and 6.8 percent for 8th-grade students.

School crime. The trend in indicators of school crime is encouraging overall; however, the level of school crime remains high, especially violent crime. In 2008, the total crime victimization rate was 47 victimizations per 1,000 students, which is substantially lower than the 121 per 1,000 reported in 1996 (Robers, Zhang, and Truman, 2010) according to the annual National Crime Victimization Survey (NCVS) conducted by the Bureau of Justice Statistics. Similarly, the percentage of students ages 12 to 18 reporting being victimized at school over a six-month period decreased from 7.6 to 4.3 percent between 1999 and 2007 based on the School Crime Supplement (SCS) to the NCVS.

Results from the SCS show that prevalence of violent crime at schools appears to be decreasing. Between 1999 and 2007, the percentage of students ages 12 to 18 reporting violent crime dropped from 2.3 to 1.6 percent. However, the percentage of public schools that recorded at least one violent incident increased between the 1999–2000 school year and 2007–08 school year, from 71.4 percent to 75.5 percent.

Relevant to some of the dynamics of school crime, in 2009, approximately 11.1 percent of students in grades 9 to 12 reported that they had been in a physical fight on school grounds during the last 12 months based on the Centers for Disease Control and Prevention's Youth Risk Behavior Surveillance System (YRBSS). In addition, in 2007, approximately 5.3 percent of students ages 12 to 18 indicated that they were afraid of attack or harm at school (SCS).

SDFSCA Program and Efforts to Address Youth ATOD Use and School Crime

To prevent and reduce youth ATOD use and school crime, program developers have created a diverse array of interventions that are intended to be implemented in elementary and secondary schools. These programs vary in their objectives, mode of delivery, intensity, and coverage. For example, while many of the programs require teachers to deliver curricula to students in classrooms, other programs are focused on schoolwide activities or on individual students who are deemed to be at "high risk" for problem behavior.

School-based prevention programs receive support from a variety of sources, including the Safe and Drug-Free Schools and Communities Act (SDFSCA) Program, which funded grants through Fiscal Year (FY) 2009.[7] As a part of the national effort to provide programming in elementary and secondary schools, this program provided funding to states to support substance abuse and violence prevention programs. Each state had both a State Education Agency (SEA) Program and a Governors' Program. Under the legislation, the SEAs allocated funds to school districts by formula (based on Title I concentration grant funding and student enrollment); the Governors' Program awarded grants to community agencies and public and private nonprofit entities. The districts and other grantees supported prevention activities at the school and community levels. While program funding was eliminated beginning in FY 2010, it did provide over $300 million in state grants in FY 2005, which covered the data collection period for this study. Funding then reached schools in practically all school districts, and the SEA program was the largest and broadest school-based substance abuse and violence prevention program nationally.

To increase the effectiveness of school-based prevention programming, the SDFSCA Program as well as other funding sources embraced the idea that such programming should be *research-based*.[8] With scarce resources available for prevention programming, only those efforts that were likely to reduce problem behavior should be supported. Along these lines, a seminal event for the SDFSCA Program in 1998 was promulgation by the U.S. Department of Education of the "Principles of Effectiveness," which included a standard that practically all efforts

[7] SDFSCA Program refers to the formula grant program that was administered by the U.S. Department of Education and implemented in school districts and communities through SEAs and governors' offices. Funding for the SDFSCA Program was eliminated beginning in FY 2010.

[8] Research-based means research has indicated that an effort achieves its intended outcomes.

supported by the program should be research-based.[9] Not surprisingly, the emphasis on research-based programming spurred researchers and program developers to examine the effectiveness of programming. It also prompted funding sources to develop and encourage the use of programs from lists of research-based programs.

The extent to which schools were actually operating research-based prevention programs and implementing them with fidelity was critical for the SDFSCA Program. They related to several aspects of prevention programming, including (a) accountability of SDFSCA Program funding recipients for operating effective programs; (b) dissemination of information on research-based programs to funding recipients; and (c) availability of affordable research-based programs that could address local problems and could be reasonably implemented with fidelity, given existing school environments. Without these ingredients, the likelihood that a school-based prevention effort will achieve its desired outcomes (e.g., reducing alcohol abuse or bullying) was viewed as being greatly diminished. Unfortunately, previous research indicated schools frequently failed to adopt research-based programs and implement them with fidelity. The remainder of the report focuses on study findings pertaining to these two factors.

Previous Research

Previous research on the prevalence and implementation fidelity of research-based prevention programs in schools has varied on the approaches taken to study these topics and on the findings. Overall, however, it suggests that concern about these issues is well justified.

Prevalence of research-based programs. Despite the importance of operating research-based prevention programming in schools, research on the prevalence of these programs is limited, and the findings from that research are inconsistent. The studies differ in a number of ways that may affect results, including (a) whether they were based on data collected at the district or school level, (b) whether they had a national or regional scope, and (c) how they defined "research-based program." In a study that examined the use of research-based programs by a purposive sample of 104 school districts in 12 states, Hallfors and Godette (2002) found that approximately 59.0 percent of the districts reported adopting at least one of six research-based programs that appeared on a list developed by Drug Strategies. These researchers also found that

[9] An exception was that local school districts could have applied to state education agencies for a waiver of the requirement of using SDFSCA funds only for research-based programs. Those non-research-based programs must have been innovative programs that demonstrated a substantial likelihood of success.

the districts were more likely to adopt non-research-based programs (e.g., DARE) than research-based programs.

Based on a survey of all school districts in North Carolina in 2001, Pankratz and Hallfors (2004) found that 72.0 percent of districts implemented research-based curricula. These curricula were included on five lists of promising and effective programs (i.e., programs appearing on lists developed by the National Institute on Drug Abuse, Drug Strategies, Centers for Disease Control and Prevention, Center for Substance Abuse Prevention, and Safe and Drug-Free Schools Program). Middle schools were more likely to use such curricula. Other factors associated with the adoption of research-based curricula were district urbanicity, prevention coordinator time available, and coordinator experience.

In a national probability sample survey of 600 districts in 1998, Hantman and Crosse (2000) collected data on activities pertaining to prevention programming in their schools during the 1997–98 school year. This study found that only 9.0 percent of districts were using substance abuse prevention programs that were consistent with existing research (e.g., curricula and instruction directed at social competencies).

In a survey of a national probability sample of 1,593 districts conducted in 1999, Rohrbach, Ringwalt, Ennett, and Vincus (2005) found that approximately 47.5 percent were implementing at least one of ten specified research-based programs (programs on the same lists used by Pankratz and Hallfors, 2004) in at least one middle school within the district. The factors associated with use of these programs included district enrollment, urbanicity, and prevention coordinator time available.

In a study of 1,905 public and private middle schools associated with the same districts in the Rohrbach et al. (2005) study, Ringwalt et al. (2002) found that approximately 34.6 percent of public middle schools were implementing one or more research-based programs. Again, the ten research-based programs were those on the lists used by Pankratz and Hallfors (2004). The factors associated with adoption of one of those programs included urbanicity, geographical region, school enrollment, and lead prevention teacher experience.

Also using data collected from schools, Gottfredson et al. (2000), and Crosse, Burr, Cantor, Hagen, and Hantman (2001) examined the quality of prevention programs based on the reports of 3,691 program providers in a national probability sample of public and private schools during the 1997–98 school year. Rather than examine the prevalence of specific programs, this

study focused on the extent to which program characteristics met criteria derived from a review of the literature. It found that 61 percent of the programs met criteria for use of best practices on content, but only 33 percent met criteria on best practices for methods.

In summary, the status of research-based prevention programming in schools is far from clear or complete. Only one study with a national scope focused on the prevalence of specific programs identified as research-based (Ringwalt et al., 2002; Rohrbach et al., 2005). At the district level, this study found that approximately 48.0 percent of districts were operating research-based programs in at least one of their schools during the late 1990s; at the school level, it found that 34.6 percent of middle schools adopted such programs.

Implementation fidelity of research-based programs. Program developers, operators, and evaluators have been paying increasing attention to the quality of program implementation. Implementation fidelity—which has frequently been discussed in terms of treatment fidelity, adherence or integrity—refers to how well a treatment or intervention is put into practice. Dusenbury, Brannigan, Falco, and Hansen (2003) define several aspects or dimensions of implementation fidelity, including (a) adherence to methods that conform to theoretical guidelines; (b) completeness or dosage of implementation; (c) quality of program delivery; (d) degree to which participants engaged; and (e) differentiation or the degree to which elements that would distinguish one program from another are present or absent. Each aspect of quality can be thought of as a continuum on which a given program can be placed. That is, some programs will perform well on a given aspect, while others will substantially miss the mark.

Implementation fidelity is important for several reasons. As mentioned, logically, one would expect that high-quality implementation would be required to achieve desired effects. Indeed, reviews of the literature and meta-analyses on preventive interventions support this relationship. For example, in a review of the research literature on the implementation of school-based substance abuse prevention programs, Dusenbury et al. (2003) report that fidelity of implementation is generally associated with improved student outcomes and with changes in factors that can mediate the effects of interventions. Based on a meta-analysis of school-based demonstration programs intended to prevent aggressive behavior, Wilson, Lipsey, and Derzon (2003) concluded that aspects of implementation fidelity accounted for a sizable amount of variability of effect sizes. In their meta-analysis on school–based substance abuse prevention programs, Tobler et al. (2000) also found aspects of implementation were among the factors most highly associated with effect sizes. Dane and Schneider (1998) point out that measuring implementation fidelity may be especially important for preventive interventions because they

are frequently offered in settings that are not conducive to high quality implementation, including limited resources and reliance on volunteer and paraprofessional service providers.

While many would agree that implementation fidelity is important, consensus is lacking on the extent to which it should be pursued rigidly or flexibly. Dane and Schneider (1998) describe three different positions on this count: (a) fidelity—interventions should be delivered as intended without adaptation; (b) pro-adaptation—providers should modify interventions to fit the settings in which they are delivered; and (c) compromise—accommodations can be made as long as the critical components of the intervention remain intact. The latter tack may be the most practical, but it does require clear specification of the critical components. In the absence of such guidance or compelling reason to do otherwise, program operators might be advised to follow closely the implementation guidelines provided by program developers.

Despite the presumed importance of high quality implementation, reviews of the research literature indicate that implementation fidelity is often lacking. These studies are most often conducted on individual prevention programs in a limited geographical area. In a review of such studies, Dusenbury et al. (2003) found that a substantial proportion of providers of five different school-based prevention programs—up to 84 percent of providers—deviated from prescribed implementation guidelines. Some of these deviations occurred even under circumstances in which researchers promoted high quality implementation.

Another set of studies has sought to assess implementation fidelity on a broader scale and across many different prevention programs. Based on the reports of providers in a national probability sample survey of schools, Gottfredson et al. (2000) found that, on average, programs intended to reduce student problem behavior met only 57 percent of the standards that the researchers judged would be required to achieve positive results. Silvia, Thorne, and Tashjian (1997) found a tremendous amount of inconsistency in how schools in 19 school districts implemented substance abuse prevention programs funded by the SDFSCA Program. Based on a national probability sample survey of providers in middle schools and on quality standards deriving from a meta-analysis on school-based substance abuse prevention, Ennett et al. (2003) found that most providers (62 percent) taught effective content, but that only a small proportion (17 percent) used effective delivery approaches.

Research on the factors associated with high-quality implementation is even rarer than that on the levels of implementation fidelity achieved. The studies on those factors suggest

that the following ones are potentially important: (a) support for implementation (such as provider training, use of program manuals, and monitoring of implementation), (b) program characteristics (such as complexity), (c) provider characteristics (such as attitudes toward the value of prevention programs), and (d) organizational characteristics (such as receptivity of the host organization) (Dusenbury et al., 2003).

Purpose of the Study

As the previous section indicates, despite the importance of the issue, little is known about the prevalence and implementation fidelity of research-based prevention programs in schools. Even less is known about the extent to which such programs were funded by the SDFSCA Program. Hence, the purpose of the study was to help fill gaps in that knowledge, focusing on the prevalence and implementation fidelity of research-based prevention programs in public elementary and secondary schools nationally and those funded by the SDFSCA Program. As such, this study collected critical information about the SDFSCA program, as authorized by Part A, Subpart 1, of the *Elementary and Secondary Education Act of 1965*. Specifically, the information allows for an assessment of the overall quality of programs that were implemented by grantees and provides baseline data for performance measures that pertain to the *Government Performance and Results Act*.[10] The questions that guide the study are presented in Exhibit 1.

Exhibit 1. Study questions

1.	**What proportion of youth substance abuse and/or school crime prevention programs in the nation's schools is research-based?**
2.	**To what extent are research-based youth substance abuse and/or school crime prevention programs implemented with fidelity?**

Within the overall Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, the Prevalence Study, which began data collection in fall 2005, was designed to address the first study question with information from surveys of a national probability sample of districts and schools on the prevention programs in

[10] Although the legislation authorizes the use of SDFSCA Program funds for a broad range of activities, the study focused on programs intended to prevent youth ATOD use and school crime, which is a subset of those activities.

operation during the 2004–05 school year. These programs were compared against the list of research-based programs that was developed for the study. Because the list of research-based programs prepared for the study was intended to reflect the results of current rigorous research on programs, it excluded some programs that appear on other lists of promising and effective programs either because there was a lack of research evidence or that the program's scope was dissimilar.[11] As a result, a school might have intended to apply the "Principles of Effectiveness" standard of implementing a research-based program but did not achieve it based on the study's criteria.

The Fidelity Study, which began data collection in spring 2006, was designed to address the second study question by surveying program coordinators in the schools with research-based programs on their implementation of those programs and of research-based practices. Aspects of implementation reported were compared to implementation standards, some of which include program developers' specifications for implementation.

Organization of Report

Following the current chapter, the report is organized into five chapters. Chapter 2 summarizes the study methodology, including the sample design, measurement, data collection, data processing, and data analysis. In Chapter 3, results are presented on the prevalence of research-based programs in public elementary and secondary schools. Chapter 4 presents results on implementation fidelity of research-based programs in public elementary and secondary schools. Chapter 5 presents findings on factors associated with implementation fidelity of research-based programs in schools. In Chapter 6, the results of the study are discussed.

Several appendixes provide additional information on the study methodology and findings. Appendix A provides detailed information on the study methodology, and Appendix B summarizes the research synthesis conducted for the study. Appendix C presents findings on SDFSCA Program funding and the adoption and implementation fidelity of research-based program; it also discusses the challenges of collecting information on fiscal data.

[11] In addition, the list excludes programs that have not been adequately evaluated, including programs that are relatively more difficult to evaluate. Rather than indicating that these programs are ineffective, insufficient evidence exists on their effectiveness.

2. METHODOLOGY

To answer the study question on the prevalence and implementation fidelity of research-based prevention programs in schools, surveys were conducted of national probability samples of public elementary and secondary schools and the districts with which they were associated. The surveys, which used both mail and Web-based approaches, gathered information on prevention programs operating during the 2004–05 school year and on the factors that may be associated with the implementation fidelity of research-based programs. An extensive critical review of the research literature was conducted in order to classify prevention programs as research-based for the Prevalence Study and Fidelity Study. The remainder of this chapter summarizes the sample design, measurement, data collection and processing, analysis, and study limitations. Appendix A provides more detailed information on the study methodology; Appendix B includes a summary report on the research synthesis conducted for the study.

Sample Design

The study had two target populations of interest that correspond to the Prevalence Study and Fidelity Study. The target population of chief interest for the Prevalence Study was public schools that offered any of grades 1 through 12 and were located in the 50 states and the District of Columbia. The sample design was two stage. First 2,500 districts were selected for the sample, and then close to 6,000 schools were sampled from these districts. Schools within sampled districts were assigned to explicit strata and sampled at rates designed to achieve approximately self-weighting (equal probability) samples within strata.

The target population of chief interest for the Fidelity Study was research-based prevention programs operating in public schools that offered any of grades 1 through 12 and were located in the 50 states and the District of Columbia. The sample design had three stages, including the two stages for the Prevalence Study. For schools with multiple programs, the third stage entailed sometimes sampling programs to help ease respondent burden.

The 2003–04 National Center for Education Statistics (NCES) Common Core of Data (CCD) Public Elementary and Secondary School Universe and Local Education Agency Universe files were the sources of the school and district sampling frames for the Prevalence Study. However, rather than starting directly with these CCD files, the study used the National

Assessment of Educational Progress (NAEP) 2004 national sample frame in order to take advantage of the many edits already undertaken to the CCD files for that frame (e.g., administrative school districts appearing on the school frame had been eliminated). In addition to the types of schools already eliminated from the NAEP sample frame, others ineligible for the Prevalence Study (e.g., state-run schools) were eliminated as part of the establishment of the frame of schools.

In the first stage of sample selection, 2,500 public school districts were selected after sorting the sample frame of districts by Census region, Metropolitan Statistical Area (MSA) status, and enrollment.[12] This sort produced an implicit stratification of districts, helping to achieve approximately proportionate representation across these characteristics. Sampling was done with probability proportional to a composite measure of size, which was a weighted count of the number of schools in the district. The choice of the weights was intended to help produce a self-weighting sample of schools within each of 27 strata that were defined by instructional level (elementary, middle, and high school), metropolitan status (center city, other urban and suburban, and rural), and percent minority (0–10 percent, 11–60 percent, and more than 60 percent).

In the second stage, the schools within the 2,500 sampled districts were stratified by a cross-classification of the same three variables used for the district stratification, and they were sampled at a rate that is conditional on the district's probability of selection. The target sample of about 6,000 schools was allocated across the 27 strata to achieve targeted levels of precision for estimates of each of the three separate categories associated with the three main school stratification variables (instructional level, metropolitan status, and percent minority).

In the third stage, to help reduce respondent burden, programs included in the Fidelity Study were subsampled. Subsampling was undertaken so that the person most knowledgeable about the research-based programs within a school was asked to report on no more than two programs, thus limiting respondent burden. As part of this subsampling, an effort was made to select "high importance" programs (one of four programs where sample yields were expected to be low) at a higher probability of selection than the other programs.

[12] Districts with at least one eligible school were eligible for the district sample.

Measurement

This section covers the data collection instruments used in the study, efforts to identify Research-Based Programs and practices, and development of implementation measures and standards.[13] The identification efforts are discussed first, because they provide important information for the development of the instruments.

Identification of research-based programs and practices. One major component of this study was the review of the research literature to identify prevention programs and practices that have been deemed effective. An initial review of the literature indicated that identifying research-based practices and specific named programs required separate approaches. To identify practices, the study drew on several meta-analyses that provided quantitative results, across many studies, on the effectiveness of practices and general program types. To identify specific programs, study staff identified and reviewed over 2,000 individual publications on programs that were judged to be effective by external sources.

Identifying *practices* that consistently yield sizable positive effects on behavior problems entailed identifying and reviewing meta-analyses on the prevention of youth ATOD use and school crime. The search covered literature published through 2004; it yielded four meta-analyses that focused on efforts to prevent or reduce problem behavior among youth. Based on those meta-analyses, a database of the reported effect sizes was developed. This database consists of over 200 effect sizes, organized by type of outcome and by different practices and general program types. The findings for practices and general program types were reviewed to flag those that had consistently high effect sizes across diverse types of outcomes.

The results indicated that seven general types of practices were consistently related to program effectiveness, of which two were delivery-related (e.g., type of program leader or facilitator) and five were content-related (e.g., containing cognitive behavioral, behavioral modeling, or behavioral modification elements). Unfortunately, the practices are so general that they lack meaning outside of the context of specific programs. For example, having a clinician lead an intervention seems irrelevant if the specific program is ineffective; also, incorporating one or more of the content-related practices, such as using cognitive behavioral approaches, is insufficient alone to achieve program effects. For that reason, the decision was made to omit

[13] In this report, a "Research-Based Program" or "Program" is a specific intervention that has been demonstrated to be effective by rigorous research; and a "research-based program" or "program" is one or more implementations or deliveries of a Research-Based Program within a given school.

these practices from the Prevalence Study, but to examine them in the Fidelity Study. Hence, the Fidelity Study considered the extent to which types of research-based practices are implemented as part of the research-based programs on which the study focuses.

The Prevalence Study developed a list of Research-Based *Programs* for the study, because of concerns about the uneven quality of research on which existing lists of effective programs was based. Development of the list entailed compiling and screening existing lists of research-based prevention programs, reviewing literature on the programs that pass the screens, and making judgments on whether the programs achieve acceptable levels of effectiveness. These tasks included the following.

- **Compiling and screening lists.** To be comprehensive, the study began with a master list of effective and promising programs developed by Mihalic (2002). She compiled 12 best practices lists from federal agencies and other sources, including the Center for Substance Abuse Prevention, Office of Juvenile Justice and Delinquency Prevention, and the Center for the Study and Prevention of Violence. The master list was screened to include only programs that are concerned with school-based prevention. In addition, programs that some list developers classified as "promising" were screened out because the research on them was insufficient.

- **Reviewing literature.** Study staff gathered and reviewed research literature on the programs that survived the screening process. After conducting extensive automated searches for the published literature, staff further screened the programs on whether the studies conducted on them met methodological standards and on whether mechanisms were in place to support widespread dissemination of the programs.[14]

- **Examining program effectiveness.** For the programs with two independent studies that met the methodological criteria, an in-depth assessment of the level of effectiveness indicated by the studies was conducted. A program was recommended for the list of research-based programs if the rigorous studies conducted on it demonstrated a pattern of program-related findings on the outcomes of interest.

[14] The methodological standards encompassed construct validity (intervention as planned was similar to the program model; intervention was implemented as planned; no expectancy, novelty, or disruption effect occurred for the treatment group; and outcome was aligned with the intervention), internal validity (minimal differential attrition occurred; for quasi-experimental designs, adequate group equating procedures were used; no local history effect occurred; and other contaminants were unlikely), and external validity/generalizability of findings (sample represented the population of interest, which was English-speaking North American school-age youths at the level of risk targeted by the program).

- **Determining the final list**. Once the research-based programs had been identified, a panel of experts in the fields of youth ATOD-use prevention, school crime prevention, and research synthesis was convened to critique the research review methodology, and to examine the appropriateness of the programs selected for inclusion and suggest other programs that may have been missed through this process.

The research review assessed quantitative results that were statistically significant and based on experimental and strong quasi-experimental designs. The study identified 19 Research-Based Programs. These programs are:

- Adolescent Alcohol Prevention Trial (AAPT)/All Stars,
- Adolescent Transitions Program
- Aggression Replacement Training,
- Alcohol Misuse Prevention (AMP),
- Anger Coping Program,
- Brainpower (Attributional Intervention),
- Child Development Project,
- Earlscourt Social Skills Group Program,
- Early Risers,
- Incredible Years,
- Know Your Body,
- Life Skills Training
- Midwestern Prevention (Project Star),
- Positive Action,
- Project Alert,
- Promoting Alternative Thinking Strategies (PATHS),
- Responding in Peaceful and Positive Ways (RIPP)/Richmond Youth against Violence Project: Responding in Peaceful and Positive Ways,

- Second Step, and
- Toward No Drug Abuse (TND).

Of these 19 Programs, the Fidelity Study focused on the 10 Programs with Provider Survey responses from more than 10 schools.

In addition, to help reconcile the study list of research-based programs with another list that became available after the research review was underway (Helping America's Youth), the U.S. Department of Education requested that two programs be included with the 19 research-based programs in analyses on the prevalence of research-based programs in schools. One of these additional programs (Skills, Opportunities, and Recognition) had been excluded because it lacked implementation materials. The other program (Too Good for Drugs) might have been included among the research-based programs but its technical report was not a published document.

Additional information on the research review and the list of research-based programs for the Prevalence Study is presented in Appendix B.

Data collection instruments. The study used four different instruments, two of which were used in both the Prevalence Study and Fidelity Study. These instruments—Prevention Program Questionnaire, District Questionnaire, Provider Questionnaires, and Principal Questionnaire—were based on instruments used by the Study on School Violence and Prevention, which was conducted by Westat and Gottfredson Associates and funded by the U.S. Departments of Education and Justice. Those earlier instruments were modified for the current study to (a) reflect the specific objectives of the Prevalence Study and Fidelity Study; (b) take advantage of research recently conducted on research-based prevention programming; and (c) increase the clarity of items on the instruments. Revised drafts of each instrument were pretested with up to nine school principals, district prevention coordinators, and school program coordinators who were not selected for the study.

The Prevention Program Questionnaire, which was completed by school principals or their designees, was intended to identify the prevention programs operated by schools during the 2004–05 school year.[15] Paper and pencil and Web-based versions of this questionnaire were

[15] Besides helping to answer the first study question, this instrument enumerated the research-based programs in the sampled schools that would be the subject of the Fidelity Study.

developed. For each of 13 types of prevention programs (e.g., curriculum and instruction), respondents indicated whether they operated that type; if they did, they completed a short series of questions on each of the specific programs that they operated, including whether or not the program received funding from the SDFSCA Program (see Appendix A for definitions of each program type.)

The District Questionnaire, which was completed by the district SDFSCA Program coordinator or prevention coordinator, was intended to provide information on aspects of district activities that were potentially associated with adoption by schools of research-based programs. The instrument was in paper and pencil format. It included the following sections: (a) planning prevention programs, (b) implementing prevention programs, (c) evaluating prevention programs, (d) district prevention programming funding and other support, (e) and background information on the coordinator. Several of the questions asked separately about programs that received funding from the SDFSCA Program.

The Provider Questionnaire was mailed to persons identified on the Prevention Program Questionnaire as the contact person for a specific research-based program. Three different Provider Questionnaires were developed for this purpose, with each one corresponding to a unique program type. The questionnaires and programs were as follows: (a) prevention curriculum, instruction, or training programs; (b) programs to change or maintain the culture or climate of the school, alter or maintain expectations for student behavior, or secure commitment to norms; and (c) programs or services for families or family members. Each questionnaire included items on the following topics: (a) objectives, (b) implementation, (c) evaluation, (d) participant characteristics, (e) instructor characteristics, (f) training and technical support, (g) school environment, and (h) respondent characteristics. The items included questions that were used in measuring the implementation of programs and practices. Some items were identical across the different questionnaires, while other items were tailored to the specific program type.

The Principal Questionnaire was mailed to principals of the schools reporting at least one research-based program. The questionnaire included items on the following topics: (a) prevention programs and practices; (b) ATOD used, school crime, and disciplinary incidents; (c) school environment; and (d) school characteristics.

Development of implementation fidelity measures. The Fidelity Study developed several measures of implementation fidelity. Each measure is based on a comparison of reports from providers (captured with the Provider Questionnaires) on the implementation of their

programs against standards for implementation. These measures were selected because they reflect important aspects of implementation that have been discussed in the research literature, and they were supported by high quality data (i.e., low item nonresponse, adequate distribution of responses, and clear basis for developing standards).

Two types of measures were used, program-specific measures and generic measures. The standards for the program-specific measures were based on the program developer's specifications for individual programs (e.g., Life Skills Training) on four aspects of implementation: (a) frequency of student participation, (b) number of lessons delivered, (c) topics covered, and (d) appropriateness of risk level targeted. Frequency of participation and number of lessons delivered were intended to gauge *completeness or dosage* of implementation, whereas the aspect of topics covered was designed to tap *adherence* (see Dusenbury et al., 2003); appropriateness of risk level targeted was used to measure whether the type of prevention program (in terms of the Institute of Medicine of the National Academies categories of universal, selective, or indicated prevention strategy) matched the characteristics of the population (same level of risk for substance abuse or school crime in the general population, subgroup with a heightened level of risk, and individuals with a heightened level of risk) to which it was delivered.

To develop the standards on which the program-specific measures are based, researchers obtained and coded aspects of the program implementation manuals and other materials for each of the research-based programs on which the Fidelity Study focused. Several programs required multiple standards for a given aspect of implementation, because those programs offered multiple components or somewhat different versions corresponding to age or grade levels.

The generic measures were based on standards applied to aspects of implementation for which developer specifications were unavailable for most or all of the programs. Hence, rather than being specific to each research-based program, the generic measures used standards that were the same for each program. The generic measures are instructional strategies (e.g., use of peer instructors or facilitators), and use of rewards, recognition, and mastery assessment (e.g., rewarding groups and individuals for their performance). As such, these measures are intended to assess the *quality of program delivery* (see Dusenbury et al., 2003).

To develop the standards on which the generic measures are based, researchers used the results of the review of meta-analyses described earlier in this chaper. They also drew on the

research-based practices identified by Gottfredson et al. (2000), which resulted from a review of the literature by the principal authors of that report.

For both the program-specific and generic measures, the study's researchers established thresholds or cut points, in consultation with the Office of Safe and Drug-Free Schools, for deciding whether the implementation of a given program (e.g., implementation of Life Skills Training at Middle School X) passed or failed on a specific aspect of implementation. For example, the threshold for number of lessons delivered was 70 percent of the program developer's specification (e.g., 70 percent of 10 lessons, or 7 lessons): Implementations that met or exceeded this threshold passed, and those that were below it failed. Additionally, a program must pass on all four program-specific measures in order to meet the minimum level of overall program implementation fidelity.

Data Collection and Processing

Data collection consisted of recruiting districts and schools, contacting potential respondents within those districts and schools, and receiving responses and following up on nonresponse. Recruitment efforts began with an introductory letter to Chief State School Officers, state prevention coordinators, and district superintendents encouraging district and school participation from a senior U.S. Department of Education official. In follow-up contacts with district officials and schools about participating in the study, study staff stressed the legitimacy of the overall study and emphasized the importance of participation in it. Ninety-six percent of the sampled districts agreed to participate in the study.

Starting in late fall 2005, once a district agreed to participate in the study, study staff identified and contacted an appropriate respondent in the district and in each selected school in that district. Staff mailed cover letters and copies of District Questionnaires to approximately 2,500 districts. Because the Prevention Program Questionnaire was intended to be administered primarily via the Web, Westat sent school IDs, user names, and passwords by email to approximately 6,000 schools.

Data collections using the Provider Questionnaire and Principal Questionnaire were based on the responses to the Prevention Program Questionnaire from schools on their use of research-based programs. In spring 2006, the study mailed copies of the Provider Questionnaire and Principal Questionnaire to each school that reported operating at least one such program in

the 2004–05 school year. A separate Provider Questionnaire was sent for each specific research-based program identified.

The study staff conducted extensive follow-up for all four of the surveys. For the data collections using paper and pencil instruments (District Questionnaire, Provider Questionnaire, and Principal Questionnaire), follow-up consisted of mailing reminders and additional copies of questionnaires, as needed, and telephone prompting. For the data collection using the Web-based instrument, follow-up entailed mailing paper and pencil versions of the questionnaire to all non-responding schools, telephone prompting, and making telephone calls to complete partially completed surveys. These efforts yielded the following response rates: 91 percent for the data collection using the District Questionnaire, 86 percent for the data collection using the Prevention Program Questionnaire, 78 percent for the data collection using the Provider Questionnaire, and 70 percent for the data collection using the Principal Questionnaire.

To ensure data quality, manual editing for the paper responses was performed directly on the survey response forms. Manual edits are designed to check each document for completeness, inter-item consistency, extraneous remarks, and proper adherence to any skip instructions. Range checks also were performed at this time. Whenever possible, sources outside the survey were used to aid in checking data for accuracy and consistency. For the Web responses, edits were performed in real time by special computer software that was programmed with built-in data checks. As needed, study staff recontacted respondents to gather missing responses, and to clarify ambiguous or inconsistent responses.

Analysis

Analysis for the study was driven by the study questions. It entailed developing sample weights, constructing derived variables and recoding variables, and conducting univariate and bivariate analyses using estimates of sample variance that reflected the complex sample design.

Westat created sample weights for analytic purposes for both the national probability sample of schools and of research-based programs. For the national probability sample of schools, these weights reflect the probabilities of selection as well as adjustments for sample schools that did not participate; for the national probability sample of research-based programs, these weights reflect the probabilities of selection as well as adjustments for nonresponse on

programs. Replicate weights were also developed to estimate sample variability, permitting the computation of standard errors and confidence intervals for school-level estimates (Prevalence Study) and for program-level estimates (Fidelity Study) that appropriately reflect the complex sample design established for this study.

The analyses consisted mainly of a univariate (e.g., percentage of research-based programs passing on a given standard) and bivariate nature (e.g., percentage of research-based programs passing on a standard by the number of students enrolled in schools with those programs). As appropriate, tests of statistical significance were conducted. To reflect the complex sample design in the estimates of variance (e.g., used in the construction of confidence intervals and in statistical tests), a replication methodology was used. After establishing strata and primary sampling units for variance estimation purposes, replicate weights for specified subsamples of the full sample were developed. WesVar software was used to develop the variance estimates.

Limitations

The study has a number of limitations that derive from the research synthesis, survey data collections, and measurement of implementation fidelity. First, the research synthesis, which was the basis for identifying research-based programs, is likely to have some under-coverage of the research literature. Despite efforts to be comprehensive, the literature searches may have missed some published studies on the programs of interest. Nonetheless, several of the programs on the study list of research-based program are geared toward elementary school-age youths, for whom these precursors are most relevant. Second, the review excluded studies that were conducted outside of the United States and English-speaking Canada. These studies were excluded because their results were deemed to have limited generalizability to students and schools in the United States. Third, some reports on studies that were reviewed omitted the information required for making judgments against criteria. In such cases, the reviewers gave the benefit of the doubt to the study.

The survey-related limitations for the Prevalence Study and Fidelity Study pertain mainly to data quality. First, schools are likely to have overreported the prevention programs operating in their schools. One form of the problem was some respondents incorrectly endorsed specific named programs, including research-based programs, because they confused those programs with other similarly named programs; another form was respondents writing in the

names of programs that probably failed to meet the definition of "program" used by the study. These two forms of the problem had a largely countervailing effect on the aggregate estimates of the prevalence of research-based programs. However, due to concerns about the school-level misreporting of research-based programs, which occurred for up to 44 percent of the original sample, we present only aggregate estimates. Second, recall problems and use of relatively less knowledgeable respondents may have diminished the quality of data, especially from schools and districts that provided information late in the data collection period. This problem arose because, although the reference period was the 2004–05 school year, data were gathered through spring 2006. Third, item nonresponse in the Prevention Program Questionnaire on whether programs received funding from the SDFSCA Program was higher than desired for several programs. Across all programs, it was approximately 30 percent. Because of concerns about the validity of the findings on receipt of funding from the SDFSCA Program, we present them separately in an appendix (Appendix C) that also discusses the challenges in collecting these data from school personnel.

The Fidelity Study has additional limitations that derive from the application of program-specific standards for assessing the fidelity of program implementation. Valid measurement of program-specific fidelity of implementation required that a program developer's specification for a program be applied to provider reports on that specific program. In some cases, provider responses raised questions about whether they were indeed reporting on the correct program (e.g., as indicated on the label on the Provider Questionnaire that they completed). Study staff attempted to confirm that providers were reporting on the correct program. If confirmation indicated that the provider reported on the wrong program, the case was considered ineligible; if confirmation could not be made and the available evidence raised concerns about whether a provider reported on the correct program, the case was excluded from analyses.

A similar problem occurred for programs that had multiple components or different versions corresponding to age or grade levels. In some cases, providers reported that their school operated multiple components or versions of a given program, each of which could have had separate standards. To deal with this issue, study staff reviewed the program materials for each component or version and sought to find the standards that best represented the components or versions; the solution often involved lowering the threshold for passing on a standard, to give the

benefit of the doubt to the program.[16] If insufficient information was reported by a provider to develop a meaningful standard, the program was omitted from analyses for that standard.

These Fidelity Study limitations pertaining to the application of program-specific standards for assessing the fidelity of program implementation contributed to achieving lower than the targeted precision for estimates. As mentioned, for a substantial number of cases, providers failed to report on the correct program. Because these cases were made ineligible or otherwise excluded from analyses, they reduced the number of valid cases. In turn, that led to decreased precision of estimates, which is reflected in larger than expected standard errors and confidence intervals. Finally, the provider survey relied on global reports by a program coordinator on how specific programs were implemented during the school year. Hence, the quality of these reports varied by the extent to which the providers were in a position to observe implementation and intentionally bias their reports. Nevertheless, previous research using similar measures (i.e., Gottfredson et al., 2000) indicated that such social desirability bias is likely to be low.

[16] Lowering the thresholds for programs led to some inflation of the estimates of implementation fidelity.

3. PREVALENCE OF PREVENTION PROGRAMS

This chapter presents findings on the prevalence of school-based ATOD use and school crime prevention programs, including research-based programs.

Prevention Programs Overall

Schools nationwide are investing resources in implementing prevention programs. More than 750,000 prevention programs were used by schools during the 2004–05 school year. This translates to millions of hours of staff and student time. This section provides information on the number and types of programs used by schools; frequently reported programs are also described.

Number of programs. The mean number of prevention programs reported to have been implemented in schools during the 2004–05 school year is 9.0. Nationally, middle schools reported 10.1 programs on average. The mean number of prevention programs reported in elementary and high schools is 8.8 and 8.9, respectively.

A wide range exists in the number of prevention programs used in schools. At one end of the spectrum, 14.8 percent of schools in the nation did not report the use of any prevention programs (see Table 1). In contrast, 11.1 percent of schools reported using more than 20 different prevention programs. The greatest numbers of schools fall into the categories 1 to 5 prevention programs (30.3 percent) and 6 to 10 prevention programs (21.9 percent).

Variation also exists in the number of prevention programs in use across instructional levels. Middle schools are somewhat more likely than elementary schools and high schools to have implemented prevention programs during the study period. While 15.6 percent of elementary schools and 15.0 percent of high schools indicated they were not implementing any prevention efforts during the 2004–05 school year, only 10.5 percent of middle schools fell into this category. Middle schools are also implementing greater numbers of prevention programs than elementary schools and high schools. For example, middle schools were more likely than elementary schools and high schools to have used more than 10 prevention programs during the 2004–05 school year.

A test of statistical significance indicated that the relationship between school instructional level and the number of programs implemented was significant, indicating that the observed relationship is not due to chance.[17]

Table 1. Schools by number of prevention programs and instructional level: 2004–05

Number of prevention programs	Instructional level									
	Elementary		Middle		High		Other		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
0	15.6	7,660	10.5	1,622	15.0	2,301	23.9	767	14.8	12,350
1 to 5	31.1	15,254	28.5	4,385	29.0	4,578	32.2	1,034	30.3	25,251
6 to 10	20.7	10,155	23.2	3,576	25.4	4,000	17.4	557	21.9	18,289
11 to 15	13.1	6,422	14.9	2,299	12.9	2,029	11.4	366	13.3	11,115
16 to 20	8.6	4,233	9.5	1,467	7.7	1,215	5.6	179	8.5	7,094
More than 20	10.8	5,316	13.3	2,046	10.3	1,620	9.5	306	11.1	9,289
Total programs	100.0	49,040	100.0	15,395	100.0	15,742	100.0	3,211	100.0	83,388

NOTE: $x^2 = 42.07$, $p<.001$. Estimates are based on responses from 4,700 schools. Per NCES, schools categorized in the tables as "other" could not be easily classified as elementary, middle, or high. Findings for schools denoted as "other" are not discussed in the text.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006; U.S. Department of Education, National Center for Education Statistics, "Common Core of Data," 2003–04.

Types of programs. Schools implemented a vast array of prevention programs in the 2004–05 school year. For ease of summarization, programs in this section are grouped into 13 discrete program types, according to a typology developed by Gottfredson et al. (2000). This carefully developed typology encompasses a wide variety of programs while providing enough nuance to distinguish programs with differing missions, goals, and approaches. (See Appendix A for definitions of the 13 program types.)

The most common prevention programs in schools are curricula, instruction, or training programs. These programs are designed to teach students factual information, increase their awareness of social influences to engage in misbehavior, expand their repertoires for recognizing and appropriately responding to risky or potentially harmful situations, etc. Nearly a quarter of all schools (23.6 percent) reported using this type of program in the 2004–05 school year (see Table 2). Programs designed to improve instructional practices were the least common type of program, reported by only 3.6 percent of schools.

[17] All references to statistical significance in this chapter are based on the standard threshold of $p < .05$.

Although schools implemented a wide variety of programs overall, the survey revealed little variation in program type by instructional level. Most program types, including curricular programs, counseling or therapeutic programs, and family programs, were implemented by very similar proportions of elementary, middle, and high schools. Even where variation across instructional levels was observed, the differences appear to be slight. Tests of significance could not be conducted on the relationship between program type and instructional level.

Table 2. Prevention programs by program type and instructional level: 2004–05

	Instructional level									
	Elementary		Middle		High		Other		All schools	
Program type	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Behavioral programming	9.7	41,698	9.0	14,027	8.1	11,362	8.7	2,085	9.2	69,172
Counseling, or therapeutic	10.6	45,710	10.8	16,872	11.3	15,911	10.6	2,531	10.8	81,024
Prevention curriculum	23.7	102,190	23.5	36,653	23.5	33,018	23.8	5,704	23.6	177,565
Mentoring, or other individual attention	4.5	19,615	4.7	7,364	5.2	7,366	5.5	1,312	4.7	35,657
Recreational, or enrichment	6.6	28,359	6.6	10,250	5.6	7,944	6.0	1,439	6.4	47,993
Improvements to instructional practices	3.8	16,343	3.3	5,100	3.1	4,389	3.8	917	3.6	26,748
Improvements to classroom management	5.9	25,363	5.3	8,234	4.3	6,102	5.3	1,266	5.4	40,965
School climate programs	8.3	36,029	7.7	12,028	7.2	10,120	7.0	1,682	8.0	59,860
School or school/community intergroup relations	4.8	20,557	5.1	7,976	5.4	7,627	5.1	1,224	5.0	37,384
Youth roles in regulating student conduct	3.4	14,587	4.1	6,312	4.1	5,812	4.5	1,083	3.7	27,794
School planning structure	3.7	16,159	4.2	6,460	4.2	5,947	4.1	976	3.9	29,542
Security or surveillance	9.3	40,037	10.0	15,532	12.4	17,471	9.8	2,353	10.0	75,393
Programs for families	5.8	25,097	5.7	8,855	5.4	7,627	5.6	1,347	5.7	42,927
Total programs	100.0	431,743	100.0	155,662	100.0	140,697	100.0	23,920	100.0	752,022

NOTE: A test of significance could not be performed. Estimates are based on responses from 4,700 schools.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006; U.S. Department of Education, National Center for Education Statistics, "Common Core of Data," 2003–04.

Frequently Reported Programs

Schools reported a wide array of programs, including programs listed (or "named") on the Prevention Program Questionnaire and other, unlisted programs. The ten most frequently cited named programs represent 47.1 percent of all named programs, but a much smaller fraction of all reported prevention programs (15.6 percent) (see Table 3). Named programs comprise all

those listed on the Prevention Program Questionnaire from which respondents were directed to identify those in use in their schools during the 2004–05 school year. Using the "other specify" fields, respondents reported the use of more than 500,000 additional programs. Analysis suggests that the number of programs identified in the other specify fields is somewhat inflated; in addition, many of these "others" are activities that do not fully meet the study definition of "program."

The ten programs are fairly diverse in focus. Five are curricular programs, while the others are divided among five other program types, including counseling programs, behavioral programming, school climate programs, school intergroup relations programs, and programs designed to improve instructional practices.

Many of the most frequently reported programs are well-known interventions that have long been used in schools. The findings should be interpreted with some caution, however, because of the somewhat generic names of many top cited programs. For example, the most frequently cited program, School Safety, could have been viewed by respondents as representing a variety of school safety-related programs or initiatives. The same could be said of the School-Based Smoking Prevention Program and Proactive Classroom Management.

DARE (Drug Abuse Resistance Education) accounts for 3.6 percent of named programs or 1.2 percent of all programs. Finding so many schools using DARE is somewhat surprising considering that evaluation research has generally not found the program to be effective (e.g., GAO, 2003). The program was not among those named on the Prevalence Survey; it was among those reported in the other specify fields. Not only was DARE the most commonly reported "other," it is more frequently used by schools than 3 other programs on the list of the 10 most reported efforts.

Table 3. Most frequently reported named prevention programs: 2004–05

Program	Percent of named programs	Percent of all programs	Weighted number of programs
School Safety Program	8.3	2.7	20,665
Behaviorally Based Prevention Program	8.1	2.7	20,275
Life Skills Training	5.2	1.7	13,072
Second Step	4.5	1.5	11,179
Children of Divorce Intervention Program	3.9	1.3	9,738
Anger Coping Program	3.7	1.2	9,155
DARE	3.6	1.2	8,866
School-based Smoking Prevention Program	3.4	1.1	8,548
Bullying Prevention Program (BPP) (The Intervention Campaign Against Bully/Victim Problems)	3.3	1.1	8,310
Proactive Classroom Management	3.2	1.0	7,871
Cumulative percent	47.1	15.6	
Total number of most frequently named programs			117,679
Total number of named programs			249,987
Total number of all programs			756,375

NOTE: Estimates are based on responses from 4,726 schools.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006.

Research-Based Programs

This section explores the extent to which schools use prevention programs shown to be effective in reducing youth substance use and school crime in rigorous evaluations. As discussed in Chapter 2, research-based programs are those that were identified through an extensive research review, which yielded a total of 21 programs. All discussion of research-based programs in the current section refers to these 21 programs.

Number of programs. A key aim of the study was to measure the prevalence of research-based prevention programs in schools. The results indicate that 40.7 percent of schools reported implementing at least one research-based program during the 2004–05 school year. Schools with research-based programs tended to implement few such programs overall, with 24.2 percent of schools reporting one research-based program, and 9.4 percent reporting two research-based programs. Only 7.1 percent of schools reported three or more research-based programs. On average, schools implemented less than one (0.7) research-based program in 2004–05; as reported, the mean number of programs implemented in schools is 9.0.

Viewed as a portion of prevention programs overall, research-based programs constitute a relatively small share of prevention programming. In the 2004–05 school year, only 7.8 percent of prevention programs overall were research-based.

Types of programs. The 21 research-based prevention programs used in analysis for the Prevalence Study can be grouped into three of the 13 program types described earlier in this chapter—prevention curricula, school climate programs, and programs for families. Of the research-based programs, the overwhelming majority are prevention curricula (97.1 percent), while the remainder is evenly divided between school climate programs (1.4 percent) and programs for families (1.5 percent). The distribution of research-based programs across program types was similar for each of the instructional levels.

Frequently reported programs. Among research-based programs, nine of the ten most frequently reported programs are curricular, while one (Know Your Body) utilizes behavioral programming (see Table 4). Two of the most frequently reported research-based programs, Life Skills Training and Second Step, were also among the most frequently reported programs overall. These programs target a variety of problem behaviors, including ATOD use and violent or aggressive behavior. The programs are also designed for a range of ages and grade levels, from preschool-age children to high school students. As discussed previously, some of these program names are fairly generic sounding, which could contribute to over-reporting.

Table 4. Most frequently reported research-based programs: 2004–05

Program	Percent of research-based programs	Percent of all programs	Weighted number of programs
Life Skills Training	22.2	1.7	13,072
Second Step	19.0	1.5	11,179
Anger Coping Program	15.6	1.2	9,155
Alcohol Misuse Prevention	9.5	0.7	5,587
Too Good For Drugs TGFD	6.7	0.5	3,920
Project ALERT	5.7	0.4	3,349
Know Your Body	5.2	0.4	3,032
Aggression Replacement Training	3.6	0.3	2,104
Positive Action	3.0	0.2	1,779
Adolescent Alcohol Prevention Trial AAPT/All Stars	2.1	0.2	1,249
Cumulative percent	92.6	7.2	
Total number of most frequently reported research-based programs			54,426
Total number of research-based programs			58,790
Total number of all programs			756,375

NOTE: These estimates are based on responses from 4,726 schools.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006.

4. IMPLEMENTATION FIDELITY OF RESEARCH–BASED PREVENTION PROGRAMS

In this chapter, results are reported on the implementation fidelity of research-based prevention programs. The discussion begins with a description of the research-based programs included in the Fidelity Study analysis and of the program-specific and generic measures used to assess program implementation fidelity. In the sections that follow, the achievement of programs on program-specific, generic, and combined multiple standards for program implementation fidelity are reported.

Research-Based Programs Included in Fidelity Analyses

The Fidelity Study focuses on assessing the implementation fidelity of research-based prevention programs provided in public schools during the 2004–05 school year. Most of the results are presented in terms of the estimated proportion and number of those programs. The study focused on curriculum programs so that the remainder of this report refers to curriculum programs rather than to programs more generally. [18]

The Fidelity Study examines the implementation of 10 research-based programs included in the Prevalence Study. Not every program identified in the research synthesis was included in the Fidelity Study. In particular, nine programs had too few schools implementing the program (fewer than 10 schools) for a reliable analysis. The study also found, based on more detailed information gathered for the Fidelity Study, that some research-based programs were misreported either because the respondent reported on a program other than the program of interest on the questionnaire, or the program of interest identified was not implemented during the 2004–05 school year. These programs had been included in the estimate of research-based programs but were subsequently taken out of the Fidelity Study in order to reflect what was actually implemented in schools. (As indicated in Chapter 2, the ineligibility of programs is likely to have had a minimal effect on the estimate of research-based programs, due to a countervailing effect of overreporting the names of programs overall.)

[18] Curriculum programs involve the provision of training or instruction to students. Results from Phase 1 indicate that they account for approximately 97 percent of the research-based prevention programs implemented in schools.

Among the 10 research-based curriculum programs included in the Fidelity Study analyses, Life Skills Training, Project Alert, and Second Step were the most frequently implemented Programs during the 2004–05 school year, accounting for 25.2, 10.8, and 39.4 percent of all programs, respectively (see Table 5). The remaining seven curriculum Programs (i.e., Aggression Replacement Training, Alcohol Misuse Prevention, All Stars, Anger Coping, Know Your Body, Positive Action, and Promoting Alternative Thinking Strategies) account for 24.6 percent of the research-based programs included in the Fidelity Study.

The majority (60.5 percent) of all research-based curriculum programs included in the Fidelity Study were in elementary schools. This is in contrast to the 24.6, 11.7, and 3.2 percent of research-based curriculum programs that were in middle, high and "other" schools, respectively. Second Step was the program most often found in elementary schools (53.0 percent) and other schools (21.9 percent), whereas Life Skills Training was the most prevalent program in middle schools (32.6 percent). Although 34.0 percent of programs in high schools were Life Skills Training, curriculum programs included in the "all other programs" category were more common (42.1 percent). The relationship between specific research-based curriculum programs and instructional level was statistically significant.

Table 5. Research-based curriculum programs in fidelity analyses by instructional level: 2004–05

Program	Instructional level[a]									
	Elementary		Middle		High		Other		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Life Skills Training ..	20.5	2,782	32.6	1,808	34.0	899	25.7	183	25.2	5,671
Project Alert.....	4.6	626	22.1	1,223	16.1	424	22.7	162	10.8	2,436
Second Step.....	53.0	7,209	22.2	1,231	7.8	206	29.7	212	39.4	8,858
All other programs.	21.9	2,980	23.0	1,274	42.1	1,111	21.9	156	24.6	5,521
Total.....	100.0	13,596	100.0	5,536	100.0	2,639	100.0	714	100.0	22,486

[a] Estimates are based on responses from 863 schools overall. $x^2 = 93.06$, $p < .001$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; U.S. Department of Education, National Center for Education Statistics, "Common Core of Data," 2003–04.

Program-Specific Measures and Generic Measures

To answer the second study question regarding how well research-based programs are being implemented, measures were developed against which program provider responses on program implementation could be compared. For the purposes of this study, two types of measures were developed: program-specific measures and generic measures. For both types of measures, the study's researchers established thresholds or cut points, in consultation with the Office of Safe and Drug-Free Schools, for deciding whether the implementation of a given program passed or failed on a specific aspect of implementation. For example, the threshold for number of lessons delivered was 70 percent of the program developer's specification (e.g., 70 percent of 10 lessons, or 7 lessons): Implementations that met or exceeded this threshold passed, and those that were below it failed.

Program-specific fidelity standards were based on criteria established by Program developers for program implementation. Two members of the evaluation team reviewed program manuals and other implementation materials for the ten curriculum programs included in the Fidelity Study analysis, to identify developer requirements or recommendations for (a) topics covered, (b) number of lessons, (c) frequency of student participation, and (d) risk level targeted (i.e., Institute of Medicine categories of universal, selective, and indicated). For programs that might use multiple manuals (i.e., because multiple grade levels were targeted), "combination" standards were established when feasible. The study standards for topics covered and number of lessons were set at 70 percent of the criteria specified by developers (see Table 6). The study standard for risk level targeted was set to be consistent with the level(s) suggested by developers, whereas the study standard for frequency of student participation was set as equal to the criteria specified by developers. A program "passed" on a given measure of implementation fidelity if it met or exceeded the standard for it.

Table 6. Program-specific and generic measures and standards by program

Program	Number of topics covered*	Number of lessons*	Frequency of student participation	Level of risk targeted	Number of instructional strategies	Number of rewards, recognition, and student mastery
Aggression Replacement Training	5	21	2 to 6 times a week	selective or indicated	4	6
Alcohol Misuse Prevention	2	3 - 6	**	universal	4	6
All Stars	2 - 7	6 - 20	2 to 6 times a week	universal or selective	4	6
Anger Coping	4	13	Once a week	indicated	4	6
Know Your Body	7 - 8	34	**	universal	4	6
Life Skills Training	5	4 - 8	Once a week	universal	4	6
Positive Action	4 - 8	13 - 97	Once a week - 2 to 6 times a week	**	4	6
Project Alert	4	2 - 8	Once a week	universal or selective	4	6
Promoting Alternative Thinking Strategies	3 - 6	7 - 39	2 to 6 times a week	universal or selective	4	6
Second Step	2 - 5	6 - 18	Once a week	universal	4	6

* Ranges are provided for programs with multiple manuals that specify different developer standards

** Developer allows too much flexibility to set a study standard, or standard is not clearly specified in implementation materials.

NOTE: Fidelity measures, with the exception of frequency of student participation and level of risk targeted, reflect approximately 70 percent of the developer specified standards.

In contrast to the program-specific measures developed for this study, the two generic fidelity measures are based on the literature regarding best practices for method of delivery (Gottfredson et al., 2000; and review of meta-analyses) rather than criteria established by Program developers. The standard for instructional strategies is based on six best practices methods listed among similar items on the Provider Questionnaire (i.e., behavioral modeling, role-playing, practice of new skills, use of cues to encourage certain behaviors, behavioral management or behavioral modification techniques, and peer teachers or leaders). Likewise, the standard for rewards, recognition, and mastery assessment is based on eight additional best practices methods also listed among similar question items on the Provider Questionnaire (i.e.,

application of rewards for individual and group achievements; student recognition for effort, improvement, successful competition against other students, and performance; assessment of student mastery and reteaching material not mastered). The study standards for instructional strategies and rewards, recognition and mastery assessment were set at 70 percent of these identified best practice methods. Again, a program "passed" on a given measure of implementation fidelity if it met or exceeded the standard for it.

Achievement of Program-Specific Standards

As discussed in the previous section, the Fidelity Study examined the implementation fidelity of research-based curriculum programs in schools, using program-specific and generic measures. This section reports results on the four individual program-specific measures: (a) topics covered, (b) number of lessons, (c) frequency of student participation, and (d) targeting on risk level.

The results indicate that over 60 percent of all programs passed on individual program-specific measures (see Table 7). The proportion of programs passing ranged from 63.4 percent for number of lessons to 92.0 percent for topics covered. The proportion passing on frequency of student participation and level of risk targeted was 64.8 percent and 88.8 percent, respectively. On any given measure, the estimated proportion passing varied by instructional level, by as much as 15 percentage points. However, with the exception of topics covered, the relationship between passing on a program-specific standard and instructional level was not statistically significant.

Table 7. Research-based curriculum programs passing on program-specific standards by instructional level: 2004–05

Passed	Instructional level									
	Elementary		Middle		High		Other		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered[a]										
Yes ...	94.1	12,577	89.8	4,921	86.3	2,276	89.0	636	92.0	20,410
No.....	5.9	784	10.2	558	13.7	363	*11.0	*78	8.0	1,783
Total .	100.0	13,361	100.0	5,479	100.0	2,639	100.0	714	100.0	22,193
Number of lessons										
Yes ...	60.2	5,980	70.3	2,946	62.1	1,049	75.0	360	63.4	10,336
No.....	39.8	3,953	29.7	1,246	37.9	641	*25.0	*120	36.6	5,959
Total .	100.0	9,932	100.0	4,193	100.0	1,690	100.0	480	100.0	16,295
Frequency of student participation										
Yes ...	62.9	7,615	69.6	3,426	65.7	1,293	60.1	405	64.8	12,739
No.....	37.1	4,483	30.4	1,499	34.3	674	39.9	269	35.2	6,925
Total .	100.0	12,099	100.0	4,925	100.0	1,966	100.0	674	100.0	19,664
Targeting on risk level										
Yes ...	90.7	11,552	87.2	4,402	84.6	2,086	78.9	553	88.8	18,593
No.....	9.3	1,180	12.8	649	15.4	379	21.1	148	11.2	2,357
Total .	100.0	12,732	100.0	5,052	100.0	2,465	100.0	702	100.0	20,950

*Unweighted n <= 5.

[a] Estimates are based on responses on 853 programs. $x^2 = 8.86$, $p < .05$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; U.S. Department of Education, National Center for Education Statistics, "Common Core of Data," 2003–04.

The Fidelity Study also examined implementation fidelity of some individual research-based curriculum programs. The results indicate that, depending on the program-specific standard, the proportion of curriculum programs passing varied by a relatively large or moderate degree (see Table 8). For example, the proportion of programs passing on topics covered ranged from 82.9 percent for Project Alert to 96.8 percent for Second Step. However, on number of lessons, it ranged from 28.7 percent passing for all other programs to 84.8 percent for Life Skills Training. None of the individual curriculum Programs appeared to consistently dominate on the quality of its implementations. The relationship between passing on a program-specific standard and individual program was statistically significant for each of the standards.

Table 8. Research-based curriculum programs passing on program-specific standards by program: 2004–05

Program	Passed					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered[a]						
Life Skills Training	89.7	5,075	10.3	584	100.0	5,659
Project Alert	82.9	2,032	17.1	419	100.0	2,451
Second Step	96.8	8,400	3.2	275	100.0	8,675
All other programs	90.9	5,069	9.1	505	100.0	5,573
Total	92.0	20,576	8.0	1,783	100.0	22,359
Number of lessons[b]						
Life Skills Training	84.8	3,568	15.2	639	100.0	4,207
Project Alert	76.3	1,668	23.7	519	100.0	2,188
Second Step	60.8	4,289	39.2	2,761	100.0	7,050
All other programs	28.7	861	71.3	2,142	100.0	3,003
Total	63.1	10,387	36.9	6,061	100.0	16,449
Frequency of student participation[c]						
Life Skills Training	71.7	3,957	28.3	1,562	100.0	5,519
Project Alert	62.5	1,513	37.5	907	100.0	2,420
Second Step	67.0	5,869	33.0	2,894	100.0	8,763
All other programs	47.0	1,471	53.0	1,657	100.0	3,128
Total	64.6	12,811	35.4	7,020	100.0	19,830
Targeting on risk level[d]						
Life Skills Training	90.1	4,938	9.9	543	100.0	5,481
Project Alert	97.8	2,318	*2.2	*51	100.0	2,369
Second Step	92.2	7,897	7.8	666	100.0	8,563
All other programs	76.1	3,536	23.9	1,109	100.0	4,645
Total	88.7	18,689	11.3	2,369	100.0	21,058

* Unweighted n <= 5.

[a] Estimates are based on responses on 860 programs. $x^2 = 12.30$, $p < .01$.

[b] Estimates are based on responses on 628 programs. $x^2 = 71.97$, $p < .001$.

[c] Estimates are based on responses on 761 programs. $x^2 = 13.72$, $p < .01$.

[d] Estimates are based on responses on 801 programs. $x^2 = 33.33$, $p < .001$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

Achievement of Generic Standards

As previously noted, the Fidelity Study examined implementation fidelity of research-based curriculum programs using two generic measures as well as the four program-specific measures. This section reports results on those two generic measures of quality: instructional level, and rewards, recognition, and mastery assessment.

The results indicate that 77.3 percent of all programs passed on instructional strategies, whereas only 32.6 percent of all programs passed on rewards, recognition, and mastery assessment (see Table 9).[19] The proportion of curriculum programs that passed on instructional strategies varied significantly across instructional levels, ranging from 64.9 percent in high schools to 81.2 percent in elementary schools. In contrast, the proportion of programs that passed on rewards, recognition, and mastery assessment varied to a lesser extent across instructional levels, ranging from 30.7 percent in elementary schools to 39.3 in high schools. Although the relationship between passing on instructional strategies and instructional level was statistically significant, the relationship between passing on rewards, recognition, and mastery assessment and instructional level was not.

Table 9. Research-based curriculum programs passing on generic standards by instructional level: 2004–05

Passed	Instructional level									
	Elementary		Middle		High		Other		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Instructional strategies[a]										
Yes	81.2	10,818	73.7	4,027	64.9	1,714	78.2	558	77.3	17,117
No	18.8	2,508	26.3	1,434	35.1	925	21.8	156	22.7	5,023
Total	100.0	13,326	100.0	5,461	100.0	2,639	100.0	714	100.0	22,140
Rewards, recognition, and mastery assessment										
Yes	30.7	4,063	33.9	1,832	39.3	1,019	34.6	247	32.6	7,160
No	69.3	9,178	66.1	3,571	60.7	1,574	65.4	467	67.4	14,790
Total	100.0	13,240	100.0	5,403	100.0	2,593	100.0	714	100.0	21,950

[a] Estimates are based on responses on 851 programs. $x^2 = 12.47$, $p < .01$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; U.S. Department of Education, National Center for Education Statistics, "Common Core of Data," 2003–04.

The Fidelity Study also examined implementation fidelity of the three most prevalent curriculum programs and all other curriculum programs using generic standards. The results indicate that the proportion of individual curriculum programs passing on instructional strategies and on rewards, recognition and mastery assessment varies by program (see Table 10). For example, the proportion of programs passing on instructional strategies ranged from 71.7 percent

[19] Although the proportion of program implementation passing on rewards, recognition, and mastery assessment is considerably less than the proportion of programs passing on other measures of quality, this finding is consistent with results from the National Study of Delinquency Prevention in Schools on a similar measure (Gottfredson et al., 2000).

for Life Skills Training to 84.5 percent for Second Step. The relationship between passing on a generic standard and individual program was statistically significant for both generic standards.

Table 10. Research-based curriculum programs passing on generic standards by program: 2004–05

Program	Passed					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Instructional strategies[a]						
Life Skills Training	71.7	4,033	28.3	1,591	100.0	5,624
Project Alert	72.1	1,754	27.9	679	100.0	2,433
Second Step.....	84.5	7,332	15.5	1,343	100.0	8,675
All other programs.....	74.1	4,132	25.9	1,441	100.0	5,573
Total	77.3	17,251	22.7	5,055	100.0	22,306
Rewards, recognition, and mastery assessment[b]						
Life Skills Training	33.5	1,887	66.5	3,745	100.0	5,632
Project Alert	36.6	888	63.4	1,535	100.0	2,423
Second Step.....	26.1	2,267	73.9	6,432	100.0	8,699
All other programs.....	40.8	2,189	59.2	3,174	100.0	5,363
Total	32.7	7,231	67.3	14,886	100.0	22,116

[a] Estimates are based on responses on 858 programs. $x^2 = 11.67$, $p < .01$.

[b] Estimates are based on responses on 848 programs. $x^2 = 10.22$, $p < .05$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

Overall Implementation Fidelity

Achievement of an individual program-specific or generic standard represents a different aspect of program implementation fidelity. On average, programs passed on 3.8 of the 6 standards. The proportion of programs that passed on: zero standards was 0.8 percent, one standard was 3.7 percent, two standards was 12.1 percent, three standards was 21.3 percent, four standards was 26.9 percent, five standards was 26.7 percent, and six standards was 8.5 percent. For the purposes of this study, achievement of all four program-specific standards was the minimum threshold for adequate overall quality of program implementation. These measures represent key design features of the program, and therefore each program-specific measure must be met. The proportion of programs implemented with "adequate" overall quality was 44.3 percent (see Table 11).[20] The relationship between overall implementation fidelity and instructional level was not statistically significant.

[20] In Table 11 and subsequent tables, overall implementation fidelity is presented in two categories. This categorization was made because the estimates for some of the cells in tables are based on sample sizes that are too small to permit valid statistical tests.

Table 11. Overall implementation fidelity of research-based curriculum programs by instructional level: 2004–05

Overall implementation fidelity	Instructional level									
	Elementary		Middle		High		Other		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Inadequate[a]……..	55.7	4,758	52.8	1,876	59.4	910	64.7	311	55.7	7,855
Adequate[b]……	44.3	3,785	47.2	1,681	40.6	623	35.3	169	44.3	6,257
Total……	100.0	8,543	100.0	3,557	100.0	1,533	100.00	480	100.00	14,112

[a] Overall implementation fidelity is defined as "inadequate" if the program did not pass on all program-specific standards.

[b] Overall implementation fidelity is defined as "adequate" if the program passed on all program-specific standards.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; U.S. Department of Education, National Center for Education Statistics, "Common Core of Data," 2003–04.

When the three most prevalent research-based programs were examined separately, the results indicate that they were fairly similar on overall implementation fidelity compared to all other programs (see Table 12). For example, the proportion of programs implemented with adequate overall quality ranged from 43.5 percent for Second Step to 56.4 percent for Life Skills Training. These programs appeared to differ from those of all other programs on overall implementation fidelity. The relationship between overall implementation fidelity and individual Program was statistically significant.

Table 12. Overall implementation fidelity of research-based curriculum programs by program: 2004–05

Program	Overall implementation fidelity[a]					
	Inadequate[b]		Adequate[c]		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Life Skills Training	43.6	1,730	56.4	2,239	100.0	3,968
Project Alert	47.6	983	52.4	1,083	100.0	2,066
Second Step.....	56.5	3,781	43.5	2,917	100.0	6,698
All other programs.....	96.9	1,429	3.1	46	100.0	1,475
Total	55.8	7,923	44.2	6,285	100.0	14,208

[a] Estimates are based on responses on 539 programs. $x^2 = 47.1$, $p < .001$.

[b] Overall implementation fidelity is defined as "inadequate" if the program did not pass on all program-specific standards.

[c] Overall implementation fidelity is defined as "adequate" if the program passed on all program-specific standards.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

5. FACTORS ASSOCIATED WITH IMPLEMENTATION FIDELITY OF RESEARCH-BASED PREVENTION PROGRAMS

In this chapter, results are reported on the association between implementation fidelity of research-based curriculum programs and a variety of program, school, and district characteristics. The chapter explores implementation differences within research-based curriculum programs, particularly those research-based programs that achieved the study's implementation standards. The characteristics examined include school and district demographic characteristics, amount and quality of training received by program implementers, district monitoring and support, and state training and technical assistance for program implementation and evaluation. Because the characteristics were analyzed individually with each implementation fidelity measure (i.e., bivariate analyses were conducted), the results may reflect the overlap or correlation of characteristics with one another.

Findings from this chapter are not generalizable to the larger set of programs outside the study's list of research-based prevention programs. The study addressed a specific question about implementation fidelity and was intended to assess implementation fidelity in research-based programs only. This part of the study focused on programs that are expected to be effective—that is, research-based prevention programs—if they were implemented well. We expected that improving outcomes with school-based programs required (a) adoption of research-based programs, and (b) effective implementation of the programs (e.g., per the developer's specifications). Therefore, this study component assessed the extent to which both of these requirements were being met. Schools implementing research-based programs could possibly differ from schools nationally.

Table 13 summarizes findings on the association between each program-specific or generic implementation fidelity standard and these characteristics. As indicated, the quality and the amount of training received by program providers are statistically associated with a broad range of implementation standards. School characteristic and state support variables also are associated with a variety of implementation standards. No significant associations were detected between any implementation fidelity standard and the district support and monitoring factors.

In the remainder of this chapter, only tables that include significant relationships are presented. For example, results on generic standards passed are included for the school

percentage of students eligible for free or reduced-price lunches, but tables for program-specific standards and overall implementation fidelity have been omitted for that factor.

School Characteristics

Analyses were conducted on the relationship (see Table 13) between quality of program implementation and school characteristics that included school urbanicity, school enrollment, and the proportion of students eligible for free or reduced-price lunches. Data on school characteristics are drawn from the NCES Common Core of Data for the 2003–04 school year (see Appendix A for further information). This section highlights findings on the relationship between these school characteristics and quality of program implementation in research-based prevention programs.

Table 13. Summary of statistically significant relationships between implementation fidelity standards, and program, school, and district characteristics: 2004–05

Characteristic		Program-specific standards				Generic quality standards		Overall implementation fidelity
		Topics covered	Number of lessons	Frequency of student participation	Level of risk targeted	Instructional strategies	Rewards, recognition, mastery assessment	
School characteristics	School enrollment		*					*
	School urbanicity		*		*			*
	School percent eligible for free or reduced-price lunches						*	
Training received	Quality of initial training		*	*	*	*		*
	Amount of training	*	*	*	*	*		*
District characteristics	District enrollment							
	District percent eligible for free or reduced-price lunches						*	
District support/monitoring	Percent of district coordinator time							
	Years since district coordinator last attended a prevention workshop or conference							
	Extent of district monitoring							
	District evaluation of programs							
State support	State provided implementation technical assistance		*		*			
	State provided evaluation technical assistance		*					

* x^2 significant at $p < .05$.

School urbanicity. A statistically significant association was found between school urbanicity and two of the four program-specific implementation fidelity measures: number of lessons and level of risk targeted (see Table 14). In both cases, a larger proportion of programs in rural schools appeared to pass on these fidelity standards than those in suburban or urban schools. As indicated, 70.4 percent of programs in rural schools met the standard for number of lessons taught, compared with 61.0 percent and 56.5 percent of those in suburban and urban schools, respectively. Similarly, 92.3 percent of programs in rural schools met the level of risk targeted standard, compared with 83.6 percent and 89.1 percent of those in suburban and urban schools, respectively. No significant relationship was found between school urbanicity and topics covered or frequency of student participation.

Table 14. Research-based curriculum programs passing on individual program-specific standards by school urbanicity: 2004–05

	School urbanicity							
	Urban		Suburban		Rural		Total	
Passed	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered								
Yes	92.2	5,905	92.0	6,064	91.7	8,391	91.9	20,360
No	7.8	500	8.0	525	8.3	758	8.1	1,783
Total	100.0	6,405	100.0	6,589	100.0	9,149	100.0	22,143
Number of lessons[a]								
Yes	56.5	2,806	61.0	2,963	70.4	4,553	63.3	10,322
No	43.5	2,162	39.0	1,892	29.6	1,918	36.7	5,972
Total	100.0	4,968	100.0	4,855	100.0	6,471	100.0	16,294
Frequency of student participation								
Yes	67.9	3,859	63.1	3,707	63.8	5,156	64.8	12,722
No	32.1	1,821	36.9	2,166	36.2	2,930	35.2	6,917
Total	100.0	5,679	100.0	5,873	100.0	8,086	100.0	19,639
Level of risk targeted[b]								
Yes	89.1	5,319	83.6	5,054	92.3	8,199	88.9	18,572
No	10.9	653	16.4	992	7.7	683	11.1	2,328
Total	100.0	5,972	100.0	6,046	100.0	8,882	100.0	20,900

[a] Estimates are based on responses on 622 programs. $x^2 = 8.00$, $p < .05$.

[b] Estimates are based on responses on 794 programs. $x^2 = 7.77$, $p < .05$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; National Center for Education Statistics, "Common Core of Data," 2003–04.

Results indicate that school urbanicity is related to overall implementation fidelity. For example, programs in rural schools (51.3 percent) were more likely to be implemented with adequate overall quality than programs in suburban (38.1 percent) and urban schools (41.9 percent) (see Table 15). The relationship between overall implementation fidelity and school urbanicity was statistically significant.

Table 15. Overall implementation fidelity of research-based curriculum programs by school urbanicity: 2004–05

Overall implementation fidelity	School urbanicity							
	Urban		Suburban		Rural		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Inadequate[a]	58.1	2,484	61.9	2,596	48.7	2,745	55.5	7,826
Adequate[b]	41.9	1,793	38.1	1,597	51.3	2,895	44.5	6,285
Total	100.0	4,277	100.0	4,194	100.0	5,640	100.0	14,111

[a] Overall implementation fidelity is defined as "inadequate" if the program did not pass on all program-specific standards.

[b] Overall implementation fidelity is defined as "adequate" if the program passed on all program-specific standards.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

School enrollment. School enrollment is associated with programs passing on number of lessons. Results indicate that 73.9 percent of programs in smaller schools (those with fewer than 300 students) passed on the number of lessons standard, compared with 61.5 percent and 52.4 percent of programs in medium and large schools, respectively (see Table 16). No other statistically significant relationships were found for school enrollment, although the relationship between school enrollment and the level of risk targeted measure approached significance (x^2= 4.95, p=0.08).

School enrollment is also related to overall implementation fidelity in research-based prevention programs. Programs in schools with less than 300 students (53.7 percent) were more likely to be implemented with adequate overall quality than programs in schools with 300 to 999 students (43.7 percent) and 1,000 or more students (29.7 percent) (see Table 17). The relationship between overall implementation fidelity and school enrollment was statistically significant. The similar findings for school urbanicity and school enrollment may be related to rural schools tending to enroll fewer students than suburban and urban schools.

Table 16. Research-based curriculum programs passing on individual program-specific standards by school enrollment: 2004–05

Passed	School enrollment							
	Less than 300		300 to 999		1,000 or more		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered								
Yes	91.5	4,540	92.3	13,589	90.8	2,191	91.9	20,319
No	8.5	420	7.7	1,140	9.2	223	8.1	1,783
Total	100.0	4,960	100.0	14,729	100.0	2,413	100.0	22,102
Number of lessons[a]								
Yes	73.9	2,656	61.5	6,807	52.4	858	63.3	10,322
No	26.1	938	38.5	4,253	47.6	781	36.7	5,972
Total	100.0	3,595	100.0	11,060	100.0	1,639	100.0	16,294
Frequency of student participation								
Yes	66.5	2,824	64.3	8,620	63.5	1,237	64.7	12,681
No	33.5	1,421	35.7	4,785	36.5	711	35.3	6,917
Total	100.0	4,244	100.0	13,405	100.0	1,949	100.0	19,598
Level of risk targeted								
Yes	89.0	4,315	90.1	12,465	80.3	1,752	88.8	18,532
No	11.0	536	9.9	1,363	19.7	429	11.2	2,328
Total	100.0	4,851	100.0	13,829	100.0	2,180	100.0	20,860

[a] Estimates are based on responses on 622 programs. $x^2 = 8.04$, $p < .05$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; National Center for Education Statistics, "Common Core of Data," 2003–04.

Table 17. Overall implementation fidelity of research-based curriculum programs by school enrollment: 2004–05

Overall implementation fidelity	School enrollment--all schools							
	Less 300		300 to 999		1,000 or more		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Inadequate[a]	46.3	1,464	56.3	5,381	70.3	980	55.5	7,826
Adequate[b]	53.7	1,700	43.7	4,170	29.7	415	44.5	6,285
Total	100.0	3,164	100.0	9,552	100.0	1,395	100.0	14,111

[a] Overall implementation fidelity is defined as "inadequate" if the program did not pass on all program-specific standards.

[b] Overall implementation fidelity is defined as "adequate" if the program passed on all program-specific standards.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

School percent eligible for free or reduced-price lunches. Percent of students eligible for the free or reduced-price lunch program is often used as an indicator of the socioeconomic status of students within the school. The analyses revealed a statistically significant relationship between that characteristic and one of the generic quality measures: rewards, recognition, and mastery assessment (see Table 18). Programs in schools with relatively high proportions of students eligible for free or reduced-price lunches appear to be more likely to

pass than others. For example, 38.7 percent of programs in schools with over 55 percent eligible for free or reduced-price lunches passed on this measure, compared with 21.5 percent of programs in schools with fewer than 25 percent eligible for free or reduced-price lunches. No other statistically significant relationships were found for percent of students eligible for the free or reduced-price lunch program.

Table 18. Research-based curriculum programs passing on individual generic standards by school percent eligible for free or reduced-price lunches: 2004–05

Passed	School percent eligible for free or reduced-price lunches							
	25 percent or less		26 to 55 percent		More than 55 percent		Total	
	Percent	Weighted Number	Percent	Weighted Number	Percent	Weighted Number	Percent	Weighted Number
Instructional strategies								
Yes	72.0	3,524	77.5	5,584	80.5	6,209	77.3	15,318
No	28.0	1,368	22.5	1,625	19.5	1,501	22.7	4,494
Total	100.0	4,892	100.0	7,209	100.0	7,710	100.0	19,811
Rewards, recognition, and mastery assessment[a]								
Yes	21.5	1,046	31.7	2,257	38.7	2,942	31.9	6,245
No	78.5	3,812	68.3	4,869	61.3	4,660	68.1	13,342
Total	100.0	4,858	100.0	7,126	100.0	7,603	100.0	19,587

[a] Estimates are based on responses on 750 programs. $x^2 = 10.69$, $p < .01$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; National Center for Education Statistics, "Common Core of Data," 2003–04.

NOTE: Categories for percent eligible for free or reduced-price lunches were based on the distribution of schools with break points more or less evenly split into three categories that had an endpoint divisible by 5.

Training on Programs

The Provider Questionnaire included several questions about the quality of initial training received by program implementers and the amount of ongoing training. This section presents findings on the relationship between training and quality of program implementation in research-based prevention programs.

Quality of initial training provided. Quality of initial training was assessed using a composite measure of provider responses, which was grouped into three categories for analysis: no initial training or low quality, medium quality, and high quality.[21] A significant relationship

[21] The Provider Questionnaire asked respondents to rate the initial training for the program along 10 measures of training quality (see pp. 117 to 118 for information about the measures). For respondents who attended the initial training, responses were coded yes=1 and no=2. Coded "yes" and summed to create a composite score ranging from 0–10. Composite scores (ranging from 0–10) were recoded into three categories so that scores ranged from 0 to 3 = 1 (Low), 4 to 7 = 2 (Medium), and 8 to 10 = 3 (High).

was found between quality of initial training and three program-specific measures of implementation fidelity: number of lessons, frequency of student participation, and level of risk targeted (see Table 19). For each of the three measures, programs with high-quality initial training passed at higher rates than those with medium-quality training or with no or low-quality training. For example, 74.8 percent of programs with high-quality training passed on number of lessons, compared with 68.7 percent of those with medium-quality training and 54.7 percent of those with no or low-quality training.

Table 19. Research-based curriculum programs passing on individual program-specific standards by quality of initial training: 2004–05

Passed	Quality of initial training							
	No training or low quality		Medium		High		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered								
Yes	92.1	6,012	92.0	4,256	94.7	5,044	92.9	15,312
No.....	7.9	519	8.0	369	5.3	280	7.1	1,168
Total	100.0	6,531	100.0	4,625	100.0	5,324	100.0	16,480
Number of lessons[a]								
Yes	54.7	2,613	68.7	2,436	74.8	3,054	65.3	8,103
No.....	45.3	2,165	31.3	1,108	25.2	1,026	34.7	4,299
Total	100.0	4,778	100.0	3,544	100.0	4,081	100.0	12,402
Frequency of student participation[b]								
Yes	57.5	3,132	63.0	2,853	78.3	3,777	66.0	9,762
No.....	42.5	2,312	37.0	1,675	21.7	1,048	34.0	5,035
Total	100.0	5,445	100.0	4,527	100.0	4,825	100.0	14,797
Level of risk targeted[c]								
Yes	81.1	4,959	90.6	3,941	95.0	4,768	88.3	13,668
No.....	18.9	1,155	9.4	407	5.0	251	11.7	1,812
Total	100.0	6,113	100.0	4,348	100.0	5,019	100.0	15,480

[a] Estimates are based on responses on 468 programs. $x^2 = 12.18$, $p < .01$.
[b] Estimates are based on responses on 557 programs. $x^2 = 15.67$, $p < .001$.
[c] Estimates are based on responses on 584 programs. $x^2 = 13.86$, $p < .001$.
SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

A significant relationship between quality of training and the instructional strategies measure also was found. Programs with high-quality initial training passed at higher rates than those with medium-quality training or with no training or low-quality training (see Table 20): 86.7 percent of programs with higher quality training passed on the instructional strategies standard, compared with 78.5 percent of those with medium-quality training and 69.0 percent of

those with no or low-quality training. No significant relationship was found between quality of initial training and the rewards, recognition, and mastery assessment standard.

Table 20. Research-based curriculum programs passing on individual generic standards by quality of initial training: 2004–05

Passed	Quality of initial training							
	No training or low quality		Medium		High		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Instructional strategies[a]								
Yes	69.0	4,510	78.5	3,632	86.7	4,615	77.4	12,756
No.....	31.0	2,021	21.5	993	13.3	709	22.6	3,724
Total	100.0	6,531	100.0	4,625	100.0	5,324	100.0	16,480
Rewards, recognition, and mastery assessment								
Yes	31.2	2,009	33.2	1,540	35.3	1,892	33.1	5,441
No.....	68.8	4,437	66.8	3,098	64.7	3,466	66.9	11,001
Total	100.0	6,446	100.0	4,639	100.0	5,357	100.0	16,442

[a] Estimates are based on responses on 627 programs. $x^2 = 18.40$, $p < .001$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

More than half (61.4 percent) of programs that received high-quality initial training were implemented with adequate overall quality, compared with 33.3 percent of programs that received no or low-quality training (see Table 21). Conversely, programs that had no or low-quality training (38.6 percent) were less likely than those with high-quality training (61.4 percent) to be implemented with adequate overall quality.

Another way to view the results on quality of initial training is in terms of the mean number of standards passed. While the mean number of fidelity standards passed for programs with high-quality training was 4.3 standards, it was 3.9 for programs with medium-quality training and 3.5 for those with no or low-quality training.

Table 21. Overall implementation fidelity of research-based curriculum programs by quality of initial training: 2004–05

Overall implementation fidelity	Quality of initial training[a]							
	No training or low quality		Medium		High		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Inadequate[b]	66.7	2,707	56.3	1,809	38.6	1,396	54.3	5,912
Adequate[c]	33.3	1,354	43.7	1,405	61.4	2,217	45.7	4,975
Total	100.0	4,061	100.0	3,213	100.0	3,613	100.0	10,887

[a] Estimates are based on responses on 406 programs. $x^2 = 18.8$, $p < .001$.

[b] Overall implementation fidelity is defined as "inadequate" if the program did not pass on all program-specific standards.

[c] Overall implementation fidelity is defined as "adequate" if the program passed on all program-specific standards.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

Amount of training provided. Several items on the Provider Questionnaire were combined to examine the relationship between amount of training and quality of program implementation in research-based prevention programs. These items covered both the duration of initial training and the frequency of ongoing training. As with the quality of initial training received by program providers, amount of training was found to have a strong relationship with a variety of implementation fidelity measures.

For all four of the quality measures, programs with a high or medium amount of training were more likely to pass on the standard than those with a low amount of training. For example, 96.9 percent of programs with a high amount of training passed on topics covered, compared with 94.1 percent of those with a medium amount of training and 89.9 percent of those with a low amount of training (see Table 22). A similar pattern was observed for the number of lessons standard. Programs with a medium amount of training (73.3 percent) were more likely to pass on the frequency of student participation standard than those with a high amount of training (66.9 percent) or those with a low amount of training (59.4 percent). This same pattern was found for the level of risk targeted standard, though the proportion of programs passing with a high amount of training was practically equivalent to those with a medium amount.

Amount of training was also found to have a significant relationship with one of the two generic fidelity standards. Approximately 87.8 percent of programs with a high amount of training passed on the instructional strategies standard, compared with 81.1 percent of those with a medium amount of training and 74.3 percent of those with a low amount of training (see Table

23). Although a similar pattern was observed for the rewards, recognition, and mastery assessment measure, the relationship was not statistically significant.

As one might expect, the overall implementation fidelity of research-based prevention programs is related to the amount of training received by program providers. Programs with a high amount of training (62.5 percent) were more likely to be implemented with adequate overall quality than those with a low amount of training (36.1 percent) (see Table 24). Conversely, programs with a low amount of training (36.1 percent) were less likely than programs with a high amount of training (62.5 percent) to be implemented with adequate overall quality. Programs with a high amount of training passed on 4.2 standards, compared to 4.1 standards for those with a medium amount of training and 3.7 standards for those with a low amount of training.

Table 22. Research-based curriculum programs passing on individual program-specific standards by amount of training: 2004–05

Passed	Amount of training							
	Low		Medium		High		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered[a]								
Yes	89.9	11,417	94.1	5,387	96.9	3,256	92.1	20,061
No.....	10.1	1,283	5.9	336	*3.1	*105	7.9	1,722
Total	100.0	12,700	100.0	5,722	100.0	3,361	100.0	21,783
Number of lessons[b]								
Yes	58.6	5,527	64.7	2,963	77.8	1,634	62.9	10,124
No.....	41.4	3,901	35.3	1,613	22.2	467	37.1	5,981
Total	100.0	9,428	100.0	4,576	100.0	2,101	100.0	16,105
Frequency of student participation[c]								
Yes	59.4	6,545	73.3	3,918	66.9	2,033	64.4	12,495
No.....	40.6	4,470	26.7	1,426	33.1	1,005	35.6	6,901
Total	100.0	11,015	100.0	5,343	100.0	3,038	100.0	19,397
Level of risk targeted[d]								
Yes	85.8	10,249	93.0	5,092	92.5	2,839	88.7	18,180
No.....	14.2	1,700	7.0	383	7.5	232	11.3	2,315
Total	100.0	11,949	100.0	5,475	100.0	3,070	100.0	20,495

* Unweighted n <=5.

[a] Estimates are based on responses on 839 programs. $x^2 = 8.87$, $p < .01$.

[b] Estimates are based on responses on 617 programs. $x^2 = 8.79$, $p < .05$.

[c] Estimates are based on responses on 745 programs. $x^2 = 9.19$, $p < .01$.

[d] Estimates are based on responses on 781 programs. $x^2 = 9.78$, $p < .01$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

Table 23. Research-based curriculum programs passing on individual generic standards by amount of training: 2004–05

Passed	Amount of training							
	Low		Medium		High		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Instructional strategies[a]								
Yes	74.3	9,392	81.1	4,642	87.8	2,952	78.2	16,987
No.....	25.7	3,255	18.9	1,080	12.2	408	21.8	4,744
Total	100.0	12,647	100.0	5,722	100.0	3,361	100.0	21,730
Rewards, recognition, and mastery assessment								
Yes	30.1	3,767	33.6	1,958	42.2	1,363	32.9	7,088
No.....	69.9	8,744	66.4	3,861	57.8	1,863	67.1	14,469
Total	100.0	12,511	100.0	5,820	100.0	3,226	100.0	21,557

[a] Estimates are based on responses on 837 programs. $x^2 = 11.11$, $p < .01$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

Table 24. Overall implementation fidelity of research-based curriculum programs by amount of training: 2004–05

Overall implementation fidelity	Amount of training[a]							
	Low		Medium		High		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Inadequate[b].....	63.9	5,117	50.0	2,044	37.5	699	56.3	7,860
Adequate[c]	36.1	2,897	50.0	2,046	62.5	1,163	43.7	6,106
Total	100.0	8,013	100.0	4,090	100.0	1,862	100.0	13,966

[a] Estimates are based on responses on 530 programs. $x^2 = 15.3$, $p < .001$.

[b] Overall implementation fidelity is defined as "inadequate" if the program did not pass on all program-specific standards.

[c] Overall implementation fidelity is defined as "adequate" if the program passed on all program-specific standards.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007.

District Characteristics

Analyses were conducted to examine the relationship between district characteristics, including student enrollment and district percent of students eligible for free or reduced-price lunches, and quality of program implementation in research-based programs. The results indicate only one significant association: district percent of students eligible for free or reduced-price lunches with the rewards, recognition, and mastery assessment implementation fidelity standard. Approximately 37.2 percent of programs in districts with over 50 percent of students eligible for free or reduced-price lunches passed on this standard, compared with 32.9 percent of those in districts with between 31 and 50 percent of students eligible, and 22.8 percent of those in

districts with 30 or fewer percent of students eligible (see Table 25). This result is similar to the earlier finding of a relationship between the rewards, recognition, and mastery assessment standard and school percent of students eligible for free or reduced-price lunches.

Federal and State Support and Policies

The District Questionnaire included several questions on the role of the state in directing training or technical assistance toward the implementation and evaluation of prevention programs. Data on these items were included in the Fidelity Study to examine the relationship between state assistance and quality of program implementation in research-based programs.

Table 25. Research-based curriculum programs passing on individual generic standards, by district percent eligible for free or reduced-price lunches: 2004–05

Passed	District percent eligible for free or reduced-price lunches							
	30 percent or less		31 to 50 percent		More than 50 percent		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Instructional strategies								
Yes	72.0	3,984	81.7	5,233	76.9	5,789	77.1	15,006
No	28.0	1,546	18.3	1,173	23.1	1,738	22.9	4,458
Total	100.0	5,530	100.0	6,406	100.0	7,527	100.0	19,464
Rewards, recognition, and mastery assessment[a]								
Yes	22.8	1,239	32.9	2,108	37.2	2,770	31.7	6,118
No	77.2	4,205	67.1	4,292	62.8	4,685	68.3	13,182
Total	100.0	5,444	100.0	6,401	100.0	7,455	100.0	19,300

[a] Estimates are based on responses on 741 programs. $x^2 = 7.47$, $p < .05$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Provider Survey," 2007; National Center for Education Statistics, "Common Core of Data," 2003–04.

NOTE: Categories for percent eligible for free or reduced-price lunches were based on the distribution of districts with break points more or less evenly split into three categories that had an endpoint divisible by 5.

State provision of training or technical assistance on implementing programs. A significant relationship was found between state-provided training or technical assistance on implementing curriculum programs and two of the four program-specific implementation fidelity measures. Approximately 73.7 percent of programs that did not receive training or technical assistance from the state passed on the number of lessons standard, compared with 61.1 percent of those that did receive technical assistance (see Table 26). Similarly, 94.8 percent of programs that did not receive state training or technical assistance passed on the level of risk targeted standard, compared with 88.1 percent of those that did receive such assistance. No significant relationship was found between state-provided training or technical assistance and the standards

for topics covered and frequency of student participation standards, though the relationship between frequency of student participation and state-provided training or technical assistance approached significance (x^2 =3.31, p=.07).

While the mean number of standards passed was 4.0 for programs in districts that did not receive state training or technical assistance on implementing programs, it was 3.8 for programs in districts that did receive state assistance.

Table 26. Research-based curriculum programs passing on individual program-specific standards by whether state-provided training or technical assistance on implementing programs: 2004–05

Passed	State-provided training or technical assistance on implementing programs					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered						
Yes	92.3	14,292	91.4	3,956	92.1	18,248
No	7.7	1,188	8.6	373	7.9	1,561
Total	100.0	15,480	100.0	4,329	100.0	19,808
Number of lessons[a]						
Yes	61.1	6,889	73.7	2,487	64.0	9,376
No	38.9	4,386	26.3	886	36.0	5,272
Total	100.0	11,276	100.0	3,373	100.0	14,648
Frequency of student participation						
Yes	62.0	8,536	70.4	2,698	63.8	11,234
No	38.0	5,236	29.6	1,134	36.2	6,370
Total	100.0	13,772	100.0	3,832	100.0	17,604
Level of risk targeted[b]						
Yes	88.1	12,781	94.8	3,890	89.6	16,671
No	11.9	1,721	5.2	215	10.4	1,935
Total	100.0	14,502	100.0	4,104	100.0	18,606

[a] Estimates are based on responses on 565 programs. $x^2 = 5.72$, $p < .05$.

[b] Estimates are based on responses on 718 programs. $x^2 = 7.16$, $p < .01$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "District Survey," 2006, "Provider Survey," 2007.

State provision of training or technical assistance on evaluating programs. The provision of state training or technical assistance on evaluating prevention programs was also examined in relation to the implementation fidelity standards in research-based programs. A significant relationship was found between this type of assistance and the number of lessons standard. Approximately 73.8 percent of programs in districts that did not receive assistance from the state passed on the number of lessons standard, compared with 60.7 percent of those in

districts that did receive this assistance (see Table 27). Results for other program-specific and generic implementation standards were not statistically significant.

Table 27. Research-based curriculum programs passing on individual program-specific standards by whether state-provided training or technical assistance on evaluating programs: 2004–05

Passed	State-provided training or technical assistance on evaluating programs					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered						
Yes	91.8	13,651	93.1	4,491	92.1	18,142
No	8.2	1,226	6.9	335	7.9	1,561
Total	100.0	14,877	100.0	4,826	100.0	19,703
Number of lessons[a]						
Yes	60.7	6,579	73.8	2,797	64.1	9,376
No	39.3	4,257	26.2	991	35.9	5,248
Total	100.0	10,836	100.0	3,788	100.0	14,624
Frequency of student participation						
Yes	62.4	8,273	67.7	2,920	63.7	11,193
No	37.6	4,976	32.3	1,394	36.3	6,370
Total	100.0	13,249	100.0	4,314	100.0	17,563
Level of risk targeted						
Yes	88.4	12,313	93.0	4,253	89.5	16,566
No	11.6	1,616	7.0	319	10.5	1,935
Total	100.0	13,929	100.0	4,572	100.0	18,501

[a] Estimates are based on responses on 564 programs. $x^2 = 6.55$, $p < .05$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "District Survey," 2006, "Provider Survey," 2007.

6. CONCLUSIONS

In this chapter, we summarize results around the study's main questions.

Prevalence of Research-Based Programming

As discussed in Chapter 1, the Prevalence Study was primarily intended to answer the following study question.

What proportion of youth substance abuse and/or school crime prevention programs in the nation's schools is research based?

This question is important because research-based prevention programs have a greater likelihood of consistently preventing or reducing youth substance abuse and school crime. Also, the U.S. Department of Education's Safe and Drug-Free Schools and Communities Act Program mandated that practically all programs funded by the Program should be research-based.[22]

Although the Prevalence Study found that public schools were implementing a large number of prevention programs during the 2004–05 school year, only 7.8 percent of these programs were research based (as defined by the Prevalence Study). Another way to consider this finding is that the mean number of programs provided in schools is 9.0 programs, but the mean number of research-based programs in schools is 0.7. Hence, the vast majority of prevention programs in operation—over 90 percent of them—lacked adequate empirical support for their effectiveness in preventing or reducing youth substance abuse and school crime.[23] This indicates that a tremendous amount of resources (e.g., for program materials and instructor time) is being allocated to activities of unproven worth. In some cases, non-research-based programs may be supplanting research-based programs that are equivalent to them on cost and feasibility.

[22] An exception is that local school districts may apply to state education agencies for a waiver of the requirement of using SDFSCA funds only for research-based programs. Those non-research-based programs must be innovative programs that demonstrate a substantial likelihood of success.

[23] Programs that lack adequate support for their effectiveness should be distinguished from programs on which adequate support indicates they are ineffective. That is, rather than indicating that the former programs are ineffective, insufficient evidence exists on their effectiveness.

In addition, the study looked at the number of schools providing at least one research-based program during the 2004–05 school year. Although the vast majority of schools provided at least one prevention program (85.2 percent), only 40.7 of schools provided at least one research-based program. As a point of comparison, in a study that was generally similar methodologically to the Prevalence Study, Ringwalt et al. (2002) found that 34.6 percent of public middle schools nationally were implementing research-based programs in 1999; and the Prevalence Study found that 46.4 percent of public middle schools implemented research-based programs nationally during the 2004–05 school year.

Implementation Fidelity of Research-Based Programming

As discussed in Chapter 1, the Fidelity Study was primarily intended to answer the following study question.

2. To what extent are research-based youth substance abuse and/or school crime prevention programs implemented with fidelity?

This question is important because previous research indicates that well-implemented research-based curriculum programs have a greater likelihood of consistently preventing or reducing youth substance abuse and school crime.

The Fidelity Study found that less than half of the research-based curriculum programs examined in this study met minimal overall implementation fidelity standards. That is, 44.3 percent of the research-based programs were implemented with adequate overall quality, passing on all four program-specific fidelity standards used. Because approximately 7.8 percent of prevention programs offered in schools are estimated to be research-based (per results from the Prevalence Study), the estimated proportion of all curriculum programs that are research-based and well-implemented is very small: only 3.5 percent. This information suggests a tremendous amount of resources, in classroom time for prevention programming alone, is being allocated to school-based prevention efforts that either lack empirical support for their effectiveness or are implemented in ways that diminish the desired effects.

The implementation quality of research-based curriculum programs varied a fair amount depending on the specific measure of quality. On topics covered and targeting on risk

level, two of the four program-specific measures, high proportions of programs passed (92.0 percent and 88.8 percent, respectively). This indicates that program providers are performing well overall on covering the general themes or topics prescribed by program developers, and they are offering curriculum programs to the appropriate audiences (in terms of risk level). One should note, however, that the topics covered measure may be somewhat inflated because of the frequent need to adjust the standards based on developers' prescriptions, to encompass schools in which several different components or versions of a specific program were delivered (e.g., to several different grade levels); these adjustments tended to dilute the standards, allowing more programs to pass than probably would have otherwise. On somewhat related measures of content, Ennett et al. (2003) found that 62 percent of providers taught effective content and Gottfredson et al. (2000) found that 76 percent of programs passed on a similar measure.[24]

On the other two program-specific measures, number of lessons and frequency of student participation, slightly less than two-thirds of program implementation passed (63.4 percent and 64.8 percent, respectively). These results suggest that a substantial number of programs will be ineffective, even if they perform well on topics covered and serve appropriate target audiences. Without achieving the prescribed amount of program exposure, providers are depriving students of desired program outcomes (e.g., delay of onset of alcohol use). The Fidelity Study findings on number of lessons, like those on topics covered, may be somewhat inflated because of the frequent need to adjust the standards based on developers' prescriptions, to encompass schools in which several different components or versions of a specific program were delivered. Nonetheless, the findings are generally consistent with those reported by Gottfredson et al. (2000): 50 percent of programs passed on number of lessons, and 65 percent passed on frequency of participation.

The findings on the two generic measures, instructional strategies and rewards, recognition, and master assessment, are quite mixed. On the former measure, 77.3 percent of programs passed. This indicates that a substantial proportion of providers are using practices (e.g., use of peer teachers or leaders) that would be expected to enhance student achievement of intended program outcomes. The proportion of programs passing on rewards, recognition and mastery assessment (32.6 percent) is substantially lower than that found on any of the other measures of quality. Nevertheless, it is consistent with the Gottfredson et al. (2000) findings that 27 percent of programs passed on a similar measure. These findings suggest that providers could

[24] Methodological differences between these studies and the current one make comparisons of results inexact.

do much more to enhance the extent to which students benefit from the program content delivered.

Limitations Pertaining to Potential Misreporting

Because of the data limitations mentioned in Chapter 2 and other sections, caution is urged in interpreting the findings presented in this report. The most important limitation pertains to potential misreporting by providers on the research-based curriculum programs of interest, because it could mean that program developers' criteria were applied to the wrong programs. In turn, that could lead to underestimates of the fidelity of implementation of the research-based programs.

The Fidelity Study attempted to minimize the effect of misreporting on programs. For reports from providers that contained ambiguous or inconsistent information on programs, study staff recontacted the respondents to confirm their reports or examined data to identify patterns that may indicate misreporting. Based on those checks, suspect reports were excluded from the analyses. Nevertheless, the data probably still include some reports on programs other than the intended ones. The extent of that misreporting, while probably low, is unknown.

Despite this and other limitations, the results are generally consistent with those from other studies. In fact, on several measures, the Fidelity Study results indicate that implementation fidelity is somewhat higher than reported by those other studies.

REFERENCES

Crosse, S., Burr, M., Cantor, D., Hagen, C.A., and Hantman, I. (2002). *Wide scope, questionable quality: Drug and violence prevention efforts in American schools*. Washington, DC: U.S. Department of Education.

Dane, A.V., and Schneider, B.H. (1998). Program integrity in primary and early secondary education: Are implementation effects out of control? *Clinical Psychology Review,* 18(1), 23–45.

Dusenbury, L., Brannigan, R., Falco, M., and Hansen, W. (2003). A review of fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research: Theory and Practice,* 18(2), 237–256.

Ennett, S.T., Ringwalt, C.L., Thorne, J., Rohrbach, L.A., Vincus, A., Simons-Rudolph, A., and Jones, S. (2003). A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. *Prevention Science,* 4(1), 1–14.

GAO. (2003). Youth illicit drug use prevention: DARE long-term evaluations and federal efforts to identify effective program. (GAO-03-172R). Washington, DC: U.S. Government Accountability Office,

Gottfredson, G.D., Gottfredson, D.C., Czeh, E.R., Cantor, D., Crosse, S.B., and Hantman, I. (2000). *National Study of Delinquency Prevention in Schools: Final report.* Ellicott City, MD: Gottfredson Associates.

Hallfors, D., and Godette, D. (2002). Will the `Principles of Effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research,* 17(4), 461–470.

Hallfors, D., Pankratz, M., and Hartman, S. (2007). Does federal policy support the use of scientific evidence in school-based prevention programs? *Prevention Science,* 8(1), 75–81.

Hantman, I., and Crosse, S. (2000). Progress in prevention: Report on the National Study of Local Education Agency Activities under the Safe and Drug-Free Schools and Communities Act. Washington, DC: U.S. Department of Education.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., and Schulenberg, J. E. (2009). *Monitoring the Future national results on adolescent drug use: Overview of key findings, 2008* (NIH Publication No. 09-7401). Bethesda, MD: National Institute on Drug Abuse.

Mihalic, S. (2002). *Matrix of programs as identified by various federal and private agencies.* Retrieved from http://www.colorado.edu/cspv/blueprints/matrix/overview.html on Oct. 15, 2003.

Pankratz, M.M., and Hallfors, D.D. (2004). Implementing evidence-based substance use prevention curricula in North Carolina public school districts. *Journal of School Health,* 74(9), 353–358.

Ringwalt, C.L., Ennett, S., Vincus, A., Rohrbach. L.A., and Simons-Rudolph, A. (2004). Who's calling the shots? Decision-makers and the adoption of effective school-based substance use curricula. *Journal of Drug Education,* 34(1), 19–31.

Ringwalt, C.L., Ennett, S., Vincus, A., Thorne ,J., Rohrbach. L.A., and Simons-Rudolph, A. (2002). The prevalence of effective substance use prevention curricula in U.S. middle schools. *Prevention Science* 3(4), 257–265.

Robers, S., Zhang, J., and Truman, J. (2010). Indicators of School Crime and Safety: 2010 (NCES 2011-002/NCJ 230812). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Rohrbach, L.A., Ringwalt, C.L., Ennett, S.T., and Vincus, A.A. (2005). Factors associated with adoption of evidence-based substance use prevention curricula in U.S. school districts. *Health Education Research,* 20(5), 514–526

Silvia, E.S., Thorne, J., and Tashjian, C. (1997). *School-based drug prevention programs: A longitudinal study in selected school districts.* Research Triangle Park, NC: Research Triangle Institute.

Tobler, N.S.R., Roona, M., Ochshorn, P., Marshall, D.G., Streke, A.V., and Stackpole, K.M. (2000). School-based adolescent drug prevention programs: 1998 Meta-analysis. *Journal of Primary Prevention,* 20(4), 275–336.

Wilson, S.J., Lipsey, M.W., and Derzon, J.H. (2003). The effects of school-based intervention programs on aggressive behavior: a meta-analysis. *Journal of Consulting and Clinical Psychology,* 71(1), 136–149.

Appendixes A Through C

APPENDIX A. METHODOLOGY

In this appendix, more detailed information is provided on the methodology used in the Prevalence Study and Fidelity Study. This information encompasses sample design, measurement, data collection, data processing, weighting and variance estimation, and analysis. Many of the methods used in the Prevalence Study also were used in the Fidelity Study; the methods used in only one of the studies are highlighted.

This combined report includes two appendixes, in addition to Appendix A. Appendix B presents a summary report on the research synthesis that identified research-based programs and practices, and Appendix C discusses efforts to collect data on program funding from school personnel.

Sample Design

In this section, the sample design for the Prevalence Study and Fidelity Study is described, including the respondent universe and sample frame and the sample selection process for districts, schools, and programs.

Respondent universe, overview of sample selection process, and sample frame. The current section covers the respondent universe, an overview of the three-stage sample, and sample frame.

Respondent universe. The target population for the Prevalence Study was the set of public schools in the U.S with the exception of a few types of schools. The major eligibility criteria for a school to be asked to participate in the Prevalence Study were that a school be a public school with any of the grades from 1 through 12 and located within the 50 states and the District of Columbia. In addition to public schools with a grade no higher than kindergarten, other public schools excluded from the targeted population were prison schools, state-run schools, federal Department of Defense and Bureau of Indian Affairs schools, ungraded schools, special education schools, and hospital schools. Most vocational schools were also excluded. Vocational-technical schools eligible for the Prevalence Study were limited to those schools with students attending only the vocational-technical school.

The target population for the Fidelity Study was the set of research-based prevention programs (i.e., intended to prevent youth substance abuse and school crime) that are offered in public elementary and secondary schools. The major eligibility criteria for a school to be asked to participate in the Fidelity Study were that it both participated in the Prevalence Study and offered one of 19 research-based prevention programs during the 2004–05 school year. (See the Measurement section beginning on page 73 for more information on how the 19 programs were selected.) The Fidelity Study sought to collect information from school personnel on these programs.

Overview of the three-stage sample. The sample design featured a three-stage sample. In the first stage, districts were selected. In the second stage, all schools in participating districts that met the eligibility criteria for schools for the Prevalence Study based on sample frame information were assigned to one of 27 strata, and an approximately self-weighting sample of schools was selected within each stratum. In the third stage (which applied only to the Fidelity Study), in schools for which the same respondent would have had to complete more than two questionnaires on specific programs, a sample of research-based prevention programs was selected. Otherwise, all programs were selected from a school.

Sample frame. The 2002–03 National Center for Education Statistics (NCES) Common Core of Data (CCD) Public Elementary and Secondary School Universe and Agency files were the original sources for the sets of schools and districts sampled for the survey. These were the most current fully edited CCD files available when sample selection was being undertaken. However, rather than starting directly with these CCD files, the National Assessment of Educational Progress (NAEP) 2005 national sample frames of districts and of public schools, developed from these CCD files, were used to take advantage of the many edits already undertaken for those files (e.g., administrative school districts and districts without schools appearing on the school frame had been eliminated, as were vocational schools with no enrollment, and closed, ungraded, special education, hospital, and prison schools). In addition to the types of schools already eliminated from the NAEP sample frame, others ineligible for the study were eliminated as part of the creation of the file from which schools would be sampled from among the schools associated with the participating districts. These included state-run schools, federal Department of Defense and Bureau of Indian Affairs schools, schools with a grade no higher than kindergarten, and schools outside the 50 states and the District of Columbia. Still other schools were eliminated later as part of the screening of schools when contacting schools and school districts about participation in the survey. For example, the eligibility of vocational-technical schools was determined during the screening process.

Sample selection of districts. In the first stage of sample selection, 2,500 public school districts were selected from the sample frame after sorting the district records by Census region, Metropolitan Statistical Area (MSA) status, and enrollment. This sort produced an implicit stratification of districts, helping to achieve approximately proportionate representation nationally across these characteristics.

Districts with at least one eligible school were eligible for the district sample. Sample districts were selected with probability proportional to a composite measure of size. This size measure was a sum of products associated with each of the 27 sample strata. Each product was the number of schools in the district in the corresponding stratum multiplied by the targeted sampling rate for that stratum. The targeted sampling rate was computed as the ratio of the targeted sample size for the stratum to the number of schools assigned to that stratum. The sum of targeted stratum sample sizes across the 27 strata was 6,000, the overall targeted sample size. The choice of this measure of size was designed to help produce an approximately self-weighting sample of schools within each of the 27 strata (Departures from the self-weighting nature of the design are discussed in more detail in Section A.5.2.1.) Targeted sampling rates varied by strata. The measure of size MOS_j for district j was computed as follows:

$$MOS_j = \sum_{i=1}^{27} r_i n_{ij}$$

where

r_i =the sampling rate to be applied when selecting the sample of schools within stratum i, i=1, 2, …, 27

n_{ij} =the number of schools in district j assigned to stratum i

The original sample of districts was supplemented with new districts (i.e., not on the NAEP public school district frame, so having no chance of selection in the original sample) that were linked in some way to a sampled district (e.g., a new district may have split off from an originally sampled district). Such new districts uniquely linked to sampled districts were added to the sample and assigned the probability of selection of the originally sampled districts.

Sample selection of schools. This section covers the stratification of schools and sample allocation of schools across strata as well as the sample selection of schools.

Stratification of schools. Twenty-seven school strata were formed by a cross-classification of three categorical variables with three values each: instructional level (elementary, middle, high), metropolitan status (central city, other urban including suburban, rural), and percent minority (defined as percentage of students who are black or Hispanic: 0–10 percent minority, 11 to 60 percent minority, over 60 percent minority). Within a given stratum, a school was sampled with probability proportional to its district weight to help achieve an approximately self-weighting sample (an overall probability of selection for schools that is equal within strata). Stratification by metropolitan status and percent minority was used because, based on previous studies, issues related to school safety and substance abuse were expected to vary by levels of these variables. Stratification by instructional level was used because prevention programs are expected to differ appreciably by instructional level.

For the purposes of stratification, instructional level was defined using the same definition employed for the National Study of School Violence and Prevention (see Table A-1). Note that schools may contribute programs in grade-level categories other than the stratum to which they were assigned. For example, a school that has grades 8 to 12 would have been assigned to the high school stratum for sampling purposes, but a program catering to only 8th and 9th graders at the school is a middle school program. The three levels of metropolitan status were formed by collapsing categories provided by the CCD file's locale variable. Percent minority also has three levels: 0–10 percent, 11–60 percent, and 61–100 percent. Cross-classifying instructional level with metropolitan status with percent minority produces 27 strata in total (3 x 3 x 3).

Table A-1. Grade-span definitions of school classification for sampling purposes

	Highest grade											
Lowest grade	1	2	3	4	5	6	7	8	9	10	11	12
1	E	E	E	E	E	E	E	E	M	M	H	H
2		E	E	E	E	E	E	E	M	M	H	H
3			E	E	E	E	E	E	M	M	H	H
4				E	E	E	M	M	M	M	H	H
5					E	E	M	M	M	M	H	H
6						E	M	M	M	M	H	H
7							M	M	M	M	H	H
8								M	M	H	H	H
9									M	H	H	H
10										H	H	H
11											H	H
12												H

E=elementary; M=middle school; H=high school.

Sample allocation of schools across strata and the sample selection of schools. As mentioned, the schools within the 2,500 sampled districts were stratified by a cross-classification of three categories each for percent minority, metropolitan status, and institutional level and were sampled with probability proportional to the district weight to achieve an approximately self-weighting sample. The target school sample of 6,000 was allocated across the 27 strata with a goal of achieving an effective sample size of at least 400 participating schools with research-based programs for school-level estimates for each of the marginal levels of the three stratification variables. For example, the allocation algorithm called for an effective sample size of participating schools with research-based programs of at least 400 for central city schools, other urban schools, and rural schools as well as for the three categories associated with percent minority and instructional level.

The effective sample size for a particular category was determined by dividing the total sample size by the design effect associated with the variation in sampling rates resulting from pooling data across the nine strata associated with that particular category. (For example, the total sample size for central city schools was obtained by summing across the schools in the various instructional level and percent minority categories associated with central cities.) The

design effects reflected only variation in sampling rates and were computed as the ratio of the estimated variance resulting from the expected variation in rates specified for the sample design being implemented to the variance one would expect for a simple random sample of the same size. A design effect could also arise due to the clustering of sample schools within districts. However, this was expected to be small because the sample for any stratification variable category was spread out across schools in many districts with little concentration expected within individual districts. As a result, the effect of clustering of schools within districts was ignored for sample allocation purposes.

As mentioned, the minimum effective sample size targeted was 400. However, because the proportion of schools with research-based programs could vary across strata, as part of the sample allocation strategy, an attempt was made to oversample strata expected to have relatively large numbers of schools with research-based programs, such as schools in central cities or high schools. Schools with high minority populations were also targeted for oversampling in an attempt to learn more about the possible impact of such programs on subpopulations of students.

As mentioned, the focus of the Prevalence Study was to identify schools with research-based programs. A previous study (Ringwalt et al., 2002) estimated that about one-third of all public middle schools contain at least one research-based program. (No data could be found on the percentage of elementary schools and high schools that have such programs.) For the purposes of sample allocation, the yield in participating schools with research-based programs overall was assumed to be 1,600. This was intentionally slightly conservative, to account for departures from yield parameters such as the proportion of schools with research-based programs, response rates, or eligibility rates. Also, the self-weighting nature of the sampling method could result in sample sizes that departed from the target for a stratum, which was the case. In addition, some schools were expected to be added to the sample (as a result of splits, the presence of new schools, etc.) during the fielding process, using an algorithm to ensure that this was done in such a way as to preserve the probabilistic nature of their inclusion. With no such departures or inclusions, a response rate of 85 percent, an eligibility rate of 100 percent, and one in three schools having at least one research-based program, the expected yield of participating schools with research-based programs was 1,700 from the targeted sample size of 6,000 schools overall.

After sorting within each of the 27 strata on a measure of enrollment, a systematic random sample of schools was selected using district weight as a measure of size and applying

the targeted sampling rate for the stratum that was used to compute the component of the district measure of size associated with the stratum. The result of sample selection is shown in Table A-2.

Table A-2. Sample sizes of schools across the 27 strata

Stratum	Percent minority	Metro status	Grade level	Sample size
1.....	<11	Central City	Elementary	79
2.....	<11	Central City	Middle school	50
3.....	<11	Central City	High school	45
4.....	<11	Other Urban	Elementary	302
5.....	<11	Other Urban	Middle school	248
6.....	<11	Other Urban	High school	194
7.....	<11	Rural	Elementary	394
8.....	<11	Rural	Middle school	393
9.....	<11	Rural	High school	480
10.....	11 to 60	Central City	Elementary	314
11.....	11 to 60	Central City	Middle school	190
12.....	11 to 60	Central City	High school	164
13.....	11 to 60	Other Urban	Elementary	307
14.....	11 to 60	Other Urban	Middle school	245
15.....	11 to 60	Other Urban	High school	183
16.....	11 to 60	Rural	Elementary	163
17.....	11 to 60	Rural	Middle school	184
18.....	11 to 60	Rural	High school	195
19.....	60+	Central City	Elementary	602
20.....	60+	Central City	Middle school	220
21.....	60+	Central City	High school	199
22.....	60+	Other Urban	Elementary	276
23.....	60+	Other Urban	Middle school	105
24.....	60+	Other Urban	High school	86
25.....	60+	Rural	Elementary	110
26.....	60+	Rural	Middle school	56
27.....	60+	rural	High school	63
Total				5,847

As mentioned, the original sample of schools was supplemented with new schools (i.e., not on the NAEP public school frame, so having no chance of selection in the original sample) that were linked in some way to a sampled school (e.g., a new school may have split off from an originally sampled school). Such new schools uniquely linked to a sampled school were added to the sample and assigned the probability of selection of the originally sampled school.

Sample selection of programs. For the Fidelity Study, programs were selected from all 1,801 schools reporting to have implemented at least one of the 19 programs targeted. In almost all such schools, all programs were included in the study and a Provider Questionnaire

was sent for each. However, to help reduce respondent burden, programs were subsampled within 134 of the 1,801 schools. The person asked to complete the Provider Questionnaire for a particular program within a school was to be the person responsible for that program in the school. However, if any person within a school was responsible for more than two programs, an attempt was made to see if another person at the school knowledgeable about the programs offered could be asked to fill out questionnaires. If so, subsampling programs might be avoided. However, if, after such an attempt, the only person available to complete questionnaires for a set of programs was responsible for three or more such programs, then subsampling was undertaken so that the person was asked to complete questionnaires for exactly two programs.

Before subsampling a set of programs, the programs were classified as "high importance" and "other." A program of "high importance" was one of four programs (among the 19) for which the sample yields were expected to be low, so an effort was made to select such programs with a high probability of selection. If only one high importance program was among the set being subsampled, it was selected with certainty and one was selected from the remaining programs with equal probabilities. If multiple high importance programs were among those to be subsampled, a sample of size of two was selected from among all the programs eligible for subsampling with probability proportionate to size. The measures of size assigned were chosen so as to include the high importance programs at a higher rate while attempting to limit the resulting variation in sampling rates as much as seemed feasible. In this way, high importance programs would be included at a greater rate but all programs had a chance of selection. The measures of size used varied, depending on the number of high importance programs and other programs among those being subsampled.

The resulting distribution of the number of sampled programs per school is given in Table A-3.

Table A-3. Variation in sample sizes of programs per school

Number of sample programs in school	Number of schools	Percent
1	1,043	57.9
2	512	28.4
3	150	8.3
4	66	3.7
5	17	0.9
6	8	0.4
7	4	0.2
8	1	0.1
Total	1,801	100.0

Measurement

This section discusses the development of data collection instruments.[25] The development of fidelity standards is discussed in Section A.6.2.

Development of data collection instruments. Four data collection instruments were used to gather data for the study: (a) Prevention Program Questionnaire, (b) District Questionnaire, (c) Provider Questionnaire, and (d) Principal Questionnaire. Data from the Prevention Program Questionnaire and District Questionnaire were used in the Prevalence Study and Fidelity Study; data from the Provider Questionnaire and Principal Questionnaire were used in the Fidelity Study only. The instruments are described in this section.

Prevention program questionnaire. The Prevention Program Questionnaire was designed to be completed by the person within a sampled school who was most knowledgeable about the programs implemented in the school during the 2004–05 school year. Respondents were encouraged to consult with their district prevention coordinators to answer questions on funding sources for the programs. Both paper and Web versions of the instrument were developed.

The questionnaire asked about specific types of programs that included a focus on alcohol, tobacco, and other drug (ATOD) use or school crime prevention. For consistency, the following terms were defined for respondents.

- **ATOD** – Includes alcohol, tobacco, illegal drugs, inhalants, and inappropriate use of prescription and over-the-counter medications.
- **School crime** – Includes illegal, violent, or disruptive behaviors that result in damage, pain, injury, or fear or result in disruptions of the school environment. Violent behaviors include bullying, verbal aggression, physical aggression, possession or use of weapons, and sexual harassment.
- **Program** – An integrated set of activities intended to achieve one or more goals and objectives. Only programs that are supported by an implementation manual or other similar documentation should be considered.

[25] In this report, a "Research-based Program" or "Program" is a specific intervention that has been demonstrated to be effective by rigorous research; and a "research-based program" or "program" is one or more implementations or deliveries of a Research-based Program within a given school during the 2004–05 school year.

The questionnaire was divided into 14 sections. The first 13 sections corresponded to the following 13 program types:

- Behavioral programming or behavior modification programs;
- Counseling, social work, psychological, or therapeutic programs;
- Prevention curriculum, instruction, or training programs;
- Mentoring, tutoring, coaching, apprenticeship, or other programs involving individual attention;
- Recreational, enrichment, or leisure programs;
- Programs involving improvements to instructional practices;
- Programs involving improvements to classroom organization or management practices;
- Programs to change or maintain the culture or climate of the school, alter or maintain expectations for student behavior, or secure commitment to norms;
- Programs focused on intergroup relations or interaction among groups within the school or between the school and the community;
- Programs related to youth roles in regulating or responding to student conduct;
- Programs involving a school planning structure or process or a method of managing change;
- Security or surveillance programs, including programs that constitute a coordinated set of strategies and are described in an implementation manual or other similar documentation; and
- Programs or services for families or family members.

Detailed descriptions of the program types were provided before respondents were asked whether they had implemented that type of program in their schools during the 2004–05 school year. For each type of program that was used or might have been used, respondents were asked to review a list of specific programs, to (a) identify specific programs used by their school; (b) provide the names of up to two people at their school who could give additional information on the implementation of each program; and (c) indicate whether each program used received Safe and Drug-Free Schools and Communities (SDFSCA) Program funding. In the final section,

respondents were asked to describe the sources of information they used to answer the questions about SDFSCA Program funding.

Once the paper version of the Prevention Program Questionnaire was developed and tested, work began on a Web version. The Web version had the advantages of performing on-line "live" error checks and bringing responses directly into a database.

District questionnaire. The District Questionnaire was mailed to persons identified by the school superintendent's office as the most appropriate for responding to questions and district's ATOD and school crime prevention programs for the 2004–05 school year. Some of the survey questions distinguished between prevention activities funded by the SDFSCA Program and those that were supported by other funding sources. Respondents were encouraged to collaborate with other staff who might have pertinent information (e.g., budget office staff). District prevention coordinators were also notified that school respondents participating in the study were asked to contact their office regarding the funding sources for programs in their school.

Provider questionnaires. A Provider Questionnaire was mailed to persons identified on the Prevention Program Questionnaire as the contact person for a specific research-based program. Three different Provider Questionnaires were developed for this purpose, with each one corresponding to a unique program type. The questionnaires and programs were as follows: (a) prevention curriculum, instruction, or training programs; (b) programs to change or maintain the culture or climate of the school, alter or maintain expectations for student behavior, or secure commitment to norms; and (c) programs or services for families or family members. Each questionnaire included items on the following topics: (a) objectives, (b) implementation, (c) evaluation, (d) participant characteristics, (e) instructor characteristics, (f) training and technical support, (g) school environment, (h) respondent characteristics. The items included questions that were used in measuring the implementation of programs and practices. Some items were identical across the different questionnaires, while other items were tailored to the specific program type.

Principal questionnaire. A Principal Questionnaire was mailed to principals of the schools with at least one research-based program. The questionnaire included items on the following topics: (a) prevention programs and practices; (b) ATOD used, school crime, and disciplinary incidents; (c) school environment; and (d) school characteristics.

Testing of instruments. Between fall 2004 and winter 2005–2006, pretests were conducted on the survey instruments for the study. The pretest participants included up to nine schools that received SDFSCA Program funding along with the corresponding district officials for those schools. As a result of the pretest, both of the instruments were revised, as well as several aspects of the planned survey procedures.

Additional testing was conducted on the Web version of the Prevention Program Questionnaire. A series of tests were performed on the application to check for accuracy of the survey application and for logics (e.g., skip patterns and ranges). After the application was satisfactorily tested, the database was loaded and further testing was done by means of unique user logins, passwords, and other identifying information for each case. Once the developmental testing was done, the integrated Web survey system was loaded to the production server. It was placed in a protected data zone where additional tests on the production server could be done by logging into the application, using a variety of browsers and browser versions.

Data Collection

Data collection was conducted between December 2004 and March 2007. First, districts were recruited and district prevention coordinators were identified. Second, to collect data for both the Prevalence Study and the Fidelity Study, a national survey of districts using the District Questionnaire and a national survey of schools using the Prevention Program Questionnaire were conducted. Third, at a later time, to collect additional data for the Fidelity Study, a national survey of schools with research-based programs using the Provider Questionnaire and Principal Questionnaire was conducted. This section describes those efforts, as well as the response rates achieved.

Recruitment. Prior to mailing recruitment materials to district superintendents, introductory letters were sent to Chief State School Officers (CSSO) and State Prevention Coordinators, along with a list of sampled districts and schools in their state. The letter outlined the purpose of the study and the design and requested that recipients encourage their districts and schools to participate in the study.

After the sample of districts and schools was selected for the study, recruitment materials were mailed to district superintendents in December 2004. The cover letter introduced the study, requested permission to contact sampled schools, and asked superintendents to name a

district coordinator to complete the District Questionnaire. District superintendents were asked to return the District Approval/District Coordinator Designation Form in a postage-paid envelope within three weeks. Telephone follow-up for nonresponse and data clarification was initiated in February 2005 and completed in June 2005.

Included in the 2,500 sampled districts were 79 districts that required completion of special clearance procedures. These 79 special clearance districts accounted for 820 sampled schools. Each district required completion and submission of standardized forms and materials for clearance requests. Submission dates and required paperwork varied from district to district. The application process began in December 2004 and ended in February 2005. Follow-up calls were conducted from January 2005 through July 2005; they resumed in fall 2005 after school reopened. Because the study continued into another school year, some districts required a reapplication process. Seventy-seven of the districts provided approval; two districts, representing 12 sampled schools, refused participation in the study.

Survey procedures. Data collection for the Fidelity Study consisted of four separate data collections: (a) national survey of districts using the District Questionnaire; (b) national survey of schools using the Prevention Program Questionnaire; (c) national survey of program providers in schools with research-based programs using the Provider Questionnaire, and (d) national survey of principals with research-based programs using the Principal Questionnaire. Data collection for the first two surveys was conducted concurrently; at a later time, data collection for the second two surveys was conducted concurrently. Packages were mailed to district prevention coordinators and school principals in October 2005; they were mailed to the providers and principals in spring 2006. The remainder of this section describes mailout activities and preparation and data collection procedures.

Mailout activities. The mailout procedures used in each of the four data collections are described here.

Mailout for the district survey. District prevention coordinators received a cover letter, business return envelope, and District Questionnaire requesting information about district-level prevention policies and programming that may have an impact on their schools' prevention activities. These potential respondents were urged to collaborate with other district staff to ensure that information accurately reflected all prevention efforts, regardless of their funding source. District prevention coordinators were asked to return the completed survey within three weeks. Nonresponse follow-up began in November 2005 and continued through July 2006.

Mailout for the prevention program survey. The Prevention Program Questionnaire, which was intended to be administered primarily via the Web, requested information about youth ATOD and school crime prevention programs that operated in schools during the 2004–05 school year. Packets mailed to school principals included a cover letter from the U.S. Department of Education and a Web Survey Information Sheet, which provided the website address and a unique username and password for accessing the questionnaire online. Principals were informed that the survey was designed to be completed by the person most knowledgeable about youth ATOD use and school crime and violence prevention programs in their schools. Potential respondents without access to the Web were given a toll-free number to contact the help desk to request a paper version of the instrument. The letter also advised the school respondents to consult with their district prevention coordinator to answer questions on funding sources for the programs. They were encouraged to call the help desk or email a study representative if they had questions about the survey, complete the questionnaire within three weeks, and keep a copy of the completed questionnaire for their files. Potential respondents without access to the Web and those who preferred to respond by mail or fax were encouraged to contact the study help desk.

In February 2006, a second package was mailed to all nonresponding schools. Due to low response to the Web survey and numerous requests from schools for a paper version, the second mailout included a paper version of the questionnaire and a business reply envelope, along with a Web Survey Information Sheet and a revised cover letter. The revised letter emphasized that the questionnaire was not as long as it appeared, because sections would be skipped if certain programs were not implemented at the school during the reference period. Respondents were given the option of completing the paper or Web version of the questionnaire. Telephone interviewers called nonresponding schools to encourage them to complete the questionnaire and to arrange to have additional copies sent to schools via mail, email, and fax.

Mailout for the provider survey. The Provider Questionnaire was designed to be completed by the person at the sampled school who was most knowledgeable about the program implementation during the 2004–05 school year. A label was affixed to the front cover of the questionnaire. Each designated program respondent received a cover letter, a business return envelope, and a Provider Questionnaire. The letter requested that respondents provide information for the program named on the label.

Mailout for the principal survey. The principal at each sampled school received a cover letter, a business return envelope, and a Principal Questionnaire requesting information about the school climate, particularly issues related to ATOD and school crime prevention programs for the 2004–05 school year. Principals were encouraged to consult with other school staff if, for example, they were new to the school and had limited knowledge regarding the 2004–05 school year. Principals were asked to return the completed survey within three weeks. Nonresponse followup began in November 2006 and continued through March 2007.

Preparation and procedures for data collection. Before data collection began, a study help desk was established and personnel were selected and trained to staff the desk. Telephone interviewers were selected and trained for conducting non-response followup data collection activities. All staff had previous experience working with public elementary and secondary schools. The recruitment staff were interviewers with more than 10 years experience in district and school recruitment and data collection.

Comprehensive interviewer training manuals were prepared for recruitment and the surveys. These included detailed instructions for each stage of data collection, an overview of the study, step-by-step instructions for contacting districts and schools, question-by-question specifications, and suggestions for handling respondent questions. Training was conducted for each data collection task, including recruitment; district, prevention program, provider, and principal surveys; and data retrieval for resolving inconsistent and missing data.

Study help desk. The help desk provided district and school personnel with assistance on the surveys. Each survey mailout included the study toll-free number and the study email address for respondents to direct questions and report problems throughout the data collection period. The help desk was staffed by individuals who had previous help desk experience. Before data collection began, help desk staff were trained. A comprehensive help desk training manual was prepared, which included detailed instructions for handling specific problems, such as accessing the Web site, navigation of the instrument (Web or paper), questions about specific questionnaire items, addressing sampling and school reconfiguration issues, and finding the most appropriate respondents. Telephone log sheets were used for recording details about each incoming call and its resolution. The help desk staff also monitored incoming emails, and responded to questions by email or telephone. Every effort was made to respond to all questions in a timely manner.

Receipt management system. The receipt control and status monitoring systems were developed for each data collection to track the flow of processing each case in the study. At the beginning of each data collection, master files were developed that defined key respondent and status variables to monitor the flow of processing of questionnaires through various stages, including interviewing assignments; survey dispositions; and mail, fax and email transmission activities. Each system included a status reporting component which was used to generate weekly reports.

For recruitment and all but the prevention program survey, the management system was developed in Access. The receipting system for the prevention program survey was part of a Web integrated system composed of two modules: the survey receipt control module and the survey data collection module. The two modules were integrated relationally, enabling them to interact with each other.

Security for the survey processing system was maintained through the use of log-ins and passwords to access the applications. Project staff accessed the system through a series of menus that permitted update of key disposition variables and provided the user with the ability to generate regularly scheduled reports on all key collection and processing events.

Nonresponse follow-up activities. Extensive telephone follow-up was conducted to obtain completed questionnaires. Interviewers contacted schools using a carefully prepared script to verify that the questionnaire had been received and to determine its status. Interviewers made arrangements for additional survey materials to be sent to schools by mail, fax, and email. Respondents were reminded that the questionnaire was available on the Internet by accessing the study website and using the school's unique username and password.

During follow-up, more than half of the schools requested a second copy of the questionnaire. In addition to remailing questionnaires to schools, RightFax software system was used for transmitting survey materials via email and fax.

Response rates. During data collection, weekly status reports were generated. Reports provided cumulative response rates, the number of questionnaires completed (by Web, mail, or telephone), the number of sampled districts and schools that were out of the scope of the study, the number of refusals, and the number of nonrespondents.

District recruitment. Recruitment packages were mailed to 2,521 public school districts. These districts include the original 2,500 sampled districts plus additional ones identified during the fielding process that were not on the sample frame but were linked to a sampled district. Three sampled districts closed, 11 combined with other districts, and four districts were ineligible for other reasons. This left a total of 2,503 eligible districts. Fewer than 2 percent of the districts (44) formally refused to participate in the study; and another 2 percent (57) did not respond to requests to respond. The final response rate was 96 percent (2,402 responding districts). Of these, 41 percent (996) were completed by fax, 50 percent (1,205) were completed by mail, and 8 percent (201) were completed by telephone (see Table A-4).

District survey. Survey materials were mailed to 2,417 public school districts. Six districts were out of scope for the study and two refused to participate. Approximately 9 percent (224) did not respond to the survey. The final response rate was 91 percent (2,185). All but 2 percent (37) were completed by mail (see Table A-5).

Prevention program survey. Survey materials were mailed to 5,577 eligible public schools. These schools included some schools identified during the fielding process that were linked to a sampled school (e.g., a school may have split off from an originally sampled school) but with no chance of selection initially as they did not appear on the sample frame of schools. One-hundred-twelve schools closed or were reconfigured, and 45 schools were ineligible for other reasons. Among the remaining 5,420 schools, 57 schools formally refused to participate in the study, and another 618 schools did not participate for other reasons. Fewer than 2 percent (94) of schools partially completed the survey on the Web. The final response rate, with the partial completes treated as nonresponse, was 86 percent (4,651). Of those, 84 percent (3,929) were completed by Web, 15 percent (692) were completed by mail, and less than 1 percent (30) was completed by telephone and fax (see Table A-6).

Provider survey. Survey materials were mailed to 2,950 program provider respondents. One-thousand, two-hundred and ninety-seven programs were ineligible because either the respondent reported on a program other than the program of interest identified by the label on the questionnaire, or the program of interest identified was not implemented during the 2004–05 school year. Among the remaining cases, 106 were made maximum contact because staff were unable to validate eligibility based on responses provided. Eight programs were made out of scope because the school was ineligible. Nine refused to participate. The final response rate was 78 percent (1,002). All but one case were completed by mail (see Table A-7).

Principal survey. Survey materials were mailed to 1,801 public school principals. Five were out of scope for the study, and six refused to participate. Twenty-seven percent (491) did not respond to the survey. The final response rate was 70 percent (1,249). All but two cases were completed by mail (see Table A-8).

Table A-4. Response for district recruitment, by geographical region and metropolitan status

Region	MC	CF	CM	CP	IN	OS	OC	RF	Total
Midwest	15	267	371	56	0	2	2	15	728
Northeast	5	192	170	24	2	1	0	12	406
South	16	311	391	66	2	0	6	5	797
West	21	226	273	55	0	0	3	12	590
Total	57	996	1,205	201	4	3	11	44	2,521

Metropolitan status	MC	CF	CM	CP	IN	OS	OC	RF	Total
Urban	16	209	201	44	2	1	0	6	479
Suburban	20	512	611	79	1	2	1	22	1,248
Rural	21	275	393	78	1	0	10	16	794
Total	57	996	1,205	201	4	3	11	44	2,521

NOTE:

MC	-	Maximum contact (no district approval)
CF	-	Completed by fax
CM	-	Completed by mail
CP	-	Completed by phone
IN	-	Ineligible
OS	-	Out of scope–closed district
OC	-	Out of scope–combined district
RF	-	Final refusal

Table A-5. Response for district survey, by geographical region and metropolitan status

Region	NR	CF	CM	OS	RF	Total
Midwest	64	9	621	1	0	695
Northeast	37	4	343	0	2	386
South	65	20	686	3	0	774
West	58	4	498	2	0	562
Total	224	37	2,148	6	2	2,417
Metropolitan status	**NR**	**CF**	**CM**	**OS**	**RF**	**Total**
Urban	48	11	396	2	0	457
Suburban	103	14	1,084	2	2	1,205
Rural	73	12	668	2	0	755
Total	224	37	2,148	6	2	2,417

NOTE:

NR - Nonresponse, questionnaires not received
CF - Completed by fax
CM - Completed by mail
OS - Out of scope
RF - Final refusal

Table A-6. Response for prevention program survey, by geographical region and metropolitan status

Region	NR	CF	CM	CP	CW	IN	OC	OS	PC	RF	Total
Midwest	135	1	198	9	1,053	12	17	29	17	19	1,490
Northeast	114	0	139	3	526	1	3	8	15	18	827
South	201	1	192	7	1,450	16	2	28	39	10	1,946
West	168	0	163	9	900	16	3	22	23	10	1,314
Total	618	2	692	28	3,929	45	25	87	94	57	5,577

Metropolitan status	NR	CF	CM	CP	CW	IN	OC	OS	PC	RF	Total
Urban	256	1	191	9	1,070	19	0	50	33	28	1,657
Suburban	269	1	342	14	1,858	16	9	18	44	24	2,595
Rural	93	0	159	5	1,001	10	16	19	17	5	1,325
Total	618	2	692	28	3,929	45	25	87	94	57	5,577

NOTE:

NR	-	Nonresponse, questionnaires not received
CF	-	Completed by fax
CM	-	Completed by mail
CP	-	Completed by phone
CW	-	Completed by web
IN	-	Ineligible
OC	-	Out of scope, combined
OS	-	Out of scope
PC	-	Partial complete
RF	-	Final refusal

Table A-7. Response for provider survey, by geographical region and metropolitan status

Region	NR	CF	CM	OS	IN	MC	RF	Total
Midwest	82	0	257	2	346	28	2	717
Northeast	77	0	154	0	148	14	1	394
South	203	0	335	4	534	42	3	1,121
West	166	1	255	2	269	22	3	718
Total	528	1	1,001	8	1,297	106	9	2,950

Metropolitan Status	NR	CF	CM	OS	IN	MC	RF	Total
Urban	190	0	282	2	387	48	5	914
Suburban	244	1	448	3	569	26	2	1,293
Rural	94	0	271	3	341	32	2	743
Total	528	1	1,001	8	1,297	106	9	2,950

NOTE:

NR - Nonresponse, questionnaires not received
CF - Completed by fax
CM - Completed by mail
OS - Out of scope
IN - Ineligible – not program of interest or program not implemented in 2004–05
MC - Maximum contact – eligible unable to complete
RF - Final refusal

Table A-8. Response for principal survey, by geographical region and metropolitan status

Region	NR	CF	CM	OS	MC	RF	Total
Midwest	84	2	341	1	13	1	442
Northeast	73	0	168	0	5	1	247
South	193	0	433	3	22	0	651
West	141	0	305	1	10	4	461
Total	491	2	1,247	5	50	6	1,801

Metropolitan Status	NR	CF	CM	OS	MC	RF	Total
Urban	179	2	318	1	26	4	530
Suburban	219	0	589	2	15	0	825
Rural	93	0	340	2	9	2	446
Total	491	2	1,247	5	50	6	1,801

NOTE:

NR - Nonresponse, questionnaires not received
CF - Completed by fax
CP - Completed by phone
OS - Out of scope
MC - Maximum contact – eligible unable to complete
RF - Final refusal

Data Processing

This section describes the separate data processing procedures developed for district recruitment, district, school prevalence, principal, and the provider surveys. The data processing staff were trained on study-specific procedures. Separate training manuals explaining the materials and procedures to be followed were prepared for telephone interviewers and for data editors/coders. Separate training sessions were also conducted for these two groups.

Processing district approval–district coordinator designation forms. Once district approval–district coordinator designation forms were received by mail, fax, or phone, they were logged into the receipt control system. The forms were reviewed by the data editing staff for incomplete information. Editors were instructed to flag any receipted forms if the approval box was not checked or if the name and title of the person providing the approval was missing. If the name of the person providing approval to conduct the study or the designated district coordinator for the district survey was someone at the school level, a case was also flagged for follow-up calls. Interviewers verified that the designee was the most appropriate person to provide approval or to respond to a survey requesting information about districtwide prevention policies and programming. Coordinator designation information was entered into an Access database and later used for the district survey.

Processing for the district survey. The procedures used for processing the district survey encompassed data coding and editing, recoding and editing, data entry and online editing, and data editing on frequency checking.

Data coding and editing. Codebook/Edit System (COED) software was used to create an automated data source file. This was used as a data dictionary and included the data file layout; a description of each questionnaire item; and a list of valid response codes, range formats, codes for nonresponse, inapplicable responses, and defined skip patterns. The source file was used for developing the data entry program in Access. The source file was also used to create a coding and editing manual, which was in an easy-to-read format and served as the main tool for coding, editing, and processing questionnaires.

A training session was conducted by the data operations supervisor for all coders and editors, during which the entire coding manual and conventions were reviewed. During the training, coders were also provided with a list of manual logic and range checks that were designed to check each questionnaire for completeness, inter-item consistency, extraneous

remarks, and proper adherence to questionnaire skip instructions. Initially, data preparation supervisors verified 100 percent of each coder's work until the coder demonstrated proficiency. Thereafter, the percentage of work verified was reduced. When a coder or editor encountered problems that were not covered by the coding manual and manual checks, he or she was referred to a data preparation supervisor who was responsible for handling all such coding and editing decisions. When necessary, the supervisor consulted with the subject experts.

When a questionnaire was received by mail, it was stamped with the date of receipt. All questionnaires received by mail, telephone, or fax were scanned for updates and the response codes and date were logged into the receipt system. Once the survey was receipted, it was sent to the data coding and editing staff. All coding and manual editing was performed directly on the questionnaire. Coders and editors also performed manual edit and range checks which were intended to identify problems quickly. When a questionnaire had missing data or inconsistent responses, a problem sheet was written by the editor indicating each of the problems needing resolution.

Recoding and editing. A large proportion of the cases had inconsistencies between related items. The completed questionnaires also had a high level of item nonresponse. As a result, additional manual editing checks and instructions were developed for the coding and editing staff to handle and resolve these problems in a consistent and systematic way. Because of the additional checks, a fresh review and verification of all previously coded work and keyed cases was conducted.

Data entry and online editing. An Access data base was used for entering questionnaire data. When cases were ready for data entry, they were assigned a batch number. Questionnaires were keyed in batches of about 50 cases. Often, the same staff responsible for coding and editing was also used for data entry. Because some online edit checks were performed at the same time that data were entered, using staff who were already familiar with the survey edit rules was efficient. All data entry was 100 percent verified with "blind" key verification, in which the second keyer had no access to the entries made by the first keyer.

Data editing and frequency checking. In addition to the range and skip pattern edits performed during data entry, data consistency and other logic checks were performed in batch editing. This included checking for consistency between data fields. The batch editing system used software written in C language that was driven by the parameter files produced from the COED source file described earlier. The software edited batch data files produced by the Access

program and generated error reports. The reports were viewed on personal computer screens or printed for use in data follow-up efforts.

When all data were entered and batch edit checks were completed, data were merged into one data set. Unweighted frequencies were also run and checked to identify and correct additional problems. Requests were made for the identification and review of specific cases, and all necessary changes were made directly to the merged data file. This process was repeated until the data were clean and ready for analysis.

Processing for the prevention program survey. The procedures for processing the prevention program survey cover the Web survey module, data coding and editing, data entry and online editing, and data retrieval followup.

Web survey module. The data preparation process used for the school prevalence survey included a combination of Web systems and manual procedures. As was done for the district survey, a coding manual, logic, and range checks were developed.

The Web survey system is the second component of the integrated Web data collection module. The system was developed with capabilities for collecting and processing survey data received via the Web, as well as for processing survey data received via other means (mail, fax, and telephone) and entered manually. The system also facilitated administrative activities, such as producing printed questionnaires, tracking and updating the status for questionnaires completed by the Web, mail, fax, or telephone, and editing for data retrieval. The system was programmed to monitor specific reporting inconsistencies and missing data. For some types of inconsistencies and nonresponse, the Web application presented respondents and keyers with on-screen messages alerting them to the possible problem and asking them to change or confirm the entries.

Data coding and editing. When a questionnaire was received by mail, it was stamped with the date of receipt. The procedures used for processing paper questionnaires were similar to those described for the district survey. All questionnaires were scanned for updates, and the response codes and date were logged into the receipt system. Once a questionnaire was receipted, it was sent to the data coding and editing staff. All coding and manual editing was performed directly on the questionnaire. Coders and editors also performed manual edit and range checks which were intended to identify problems quickly. When a questionnaire had

missing data or inconsistent responses, a problem sheet was written by the editor indicating each of the problems needing resolution.

Data entry and online editing. Paper questionnaires were keyed by data entry staff into the Web application. From that point in time, data from both the paper and Web questionnaire followed the same process. This ensured that survey responses were subjected to the same set of built-in logics, ranges, and validation steps. To minimize respondent errors that might result from mode differences, all paper questionnaires were "double-keyed" in a simulated Web site interface and subsequent data checks for consistency were conducted.

Data retrieval followup. Follow-up was also conducted for missing or inconsistent responses. A problem sheet was prepared outlining the problem(s) for telephone follow-up. If a questionnaire completed on the Web failed one or more logic checks, or was identified as having missing data, a copy of the completed questionnaire was printed from the Web. For questionnaires received by mail, fax, or telephone, these problems were identified during coding and editing. The same telephone interviewers working on nonresponse follow-up were responsible for data retrieval follow-up, ensuring that they were thoroughly familiar with each survey item. Interviewers also received a short training on the most common data retrieval problems.

Processing for the provider and principal surveys. The procedures used for processing the principal and provider surveys were similar to those described earlier for the district survey. For both surveys, the processing encompassed data coding, editing, data entry, online editing, and frequency checks. Automated source files were created for both questionnaires using Codebook/Edit System (COED) software. The source files were then used to generate the Access data base for entering data. Data processing files followed the same procedures outlined earlier for coding the cases and preparing them for data inconsistency checks.

Once questionnaires were keyed, online edit checks were generated, and batch edits were performed to check for data consistency and other logic checks. Finally, unweighted frequencies were generated and checked to identify and correct additional problems.

Because establishing the eligibility of programs was important (i.e., to be valid, provider reports on a given program had to be compared with the program developer's specification for that program to assess the degree to which the reports were consistent with the

specification), extensive data retrieval was done for the provider survey. Cases were flagged if inconsistent information was provided about the titles and publishers for the manuals used for implementing the programs. Respondents often provided information about programs that were implemented in years other than the 2004–05 academic school year. Data retrieval was performed to ensure that information was provided for the relevant year. The status of cases was changed from complete to ineligible if data retrieval established ineligibility for either reason.

If data retrieval could not be completed and the available evidence raised concerns about whether a provider reported on the correct program, the case typically was excluded from analyses. The process for deciding on which cases to exclude involved two steps. First, cases were sorted into groups based on the amount and quality of information indicating the degree to which provider reports were consistent with the program in question (e.g., reports on the titles and publishers for manuals used for implementing the programs were compared to those specified for use by the program's developers). Second, for cases in the groups associated with the more incomplete or inconsistent information (e.g., no information on manuals and publishers), provider reports on aspects of implementation (number of lessons, level of risk, number of student participants, and grade levels targeted) were examined in detail. Two staff persons independently applied written rules for flagging cases with provider reports that were inconsistent with expectations for the correct program on two or more of the items. Their judgments were in high initial agreement; they discussed and reconciled any discrepancies. In all, 91 cases were ultimately excluded from analysis because this process raised concerns about whether a provider's reports were actually for the program in question.

Weighting and Variance Estimation

This section describes the calculation of program weights and variance estimates for use in the analysis of the Prevalence Study and Fidelity Study data. Two types of weights were created: School weights were used in analyses for the Prevalence Study; program weights were used in analyses for the Fidelity Study.

- The *school* base weight was first calculated to reflect the school's overall probability of selection, including the sampling of districts at the first stage. This base weight was then adjusted for schools whose eligibility status was unknown (where district level nonresponse was also addressed) and for nonresponse to the Prevention Program Questionnaire to help control for nonresponse bias.

- The *program* base weight was calculated to reflect the program's overall probability of selection as well as adjustments for nonresponse at the district and school levels. This base weight was then adjusted for programs whose eligibility status was unknown and for nonresponse to the Provider Questionnaire to help reduce the potential for incurring nonresponse bias.

A set of jackknife replicate weights was also created for each type of weight, to permit the calculation of standard errors that reflect the sample design.

Review of sample design. The program sample was selected in three stages. First, 2,500 districts were sampled with probability proportional to a composite measure of size, after sorting by Census region, metropolitan status, and district enrollment to create implicit strata. The composite measure of size for each district was derived from district school counts for each stratum and stratum school sampling rates. Second, schools in the sampled districts were stratified into 27 strata without regard to district membership, then sorted by enrollment and sampled with probability proportional to size within strata, with a measure of size designed to produce a self-weighting sample within strata. As expected, some sampled districts had no sampled schools. Twenty-one districts were added to the sample as a result of splits and mergers, bringing the total to 2,521, of which 170 had no sampled schools. A total of 5,897 schools were sampled, including 50 new schools that were added to the sample as a result of splits and mergers. Third, to reduce respondent burden, a subsample of programs was selected within 134 of the 1,801 schools that reported on the prevention programs they offered during the 2004–05 school year. In all, a total of 2,950 programs in 1,801 schools were included in the sample. A discussion of the sample selection of programs appears in Section A.1.4.

Weighting. Prior to analysis, data were weighted to reflect the probability of sample selection and nonresponse. This entailed developing and adjusting district and school weights; for the Fidelity Study, it also entailed developing and adjusting program weights.

District and prevention program surveys. District and school base weights were developed, and adjustments were made to those weights to reflect nonresponse, as described in the remainder of this section.

District Base Weights. The district base weight d_i for each district i among the 2,500 originally sampled districts was calculated as the inverse of the district probability of selection:

$$d_i = \frac{\sum_{i=1}^{N} mos_i}{2{,}500 * mos_i} ,$$

where N is the number of districts on the sampling frame, and mos_i is the composite measure of size for the i-th district. The probability of selection for each of 21 new districts was assigned to be that of the originally sampled district through which it came into the sample via splits or mergers of school districts. A new district could be eligible for survey participation only if it did not appear on the sample frame and hence had no chance of selection via the sample selection of districts. Districts were coded as ineligible if they merged with or were actually part of another district that had a probability of selection associated with a district record not sampled from the sample frame.

The sum of the district base weights was 14,980, representing an estimate of the public school districts eligible for sample selection for the study.

School base weights. The school base weight c_{hj} for school j associated with district i and in stratum h among the 5,847 originally sampled schools was calculated as the product of the district base weight and the inverse of the school's probability of selection within stratum h, conditional on its district being sampled:

$$c_{hij} = d_i / \min (d_i r_h, 1)$$

where r_h was the targeted sampling rate for schools in stratum h used in the development of the composite measure of size used in the sample selection of districts. The value min (d_i r_h, 1) represents the conditional probability of selection of a school found in stratum h and coming from district i. Hence, the reciprocal of this value is the corresponding component of the overall school base weight. In most cases, computing the overall school base weight c_{hij} as the product of the district base weight and the conditional school weight resulted in the value ($1/ r_h$), the targeted weight for the self-weighting design. However, in the event that d_i r_h exceeded 1, then a school was selected with certainty within a stratum, and its base weight was computed as the product of d_i and 1, so that the school base weight was also the district base weight.

The school base weights for 50 new schools added to the sample were assigned on a case-by-case basis by linking each new school to the originally sampled school it had split off from or merged with. The new school's probability of selection was assigned to that of the school through which it came into the sample. Schools were coded as ineligible if they merged with or were actually part of another school that had a probability of selection associated with a school record not sampled from the sample frame.

The school base weights sum to 87,806, representing an estimate of the total number of schools eligible for sample selection for the Prevention Program Questionnaire.

Adjustment for school nonresponse within nonparticipating districts. When a district was determined to be ineligible, the schools in the district were also treated as ineligible. However, 110 sampled schools were in districts that did not respond to the district recruitment effort, so their eligibility for the Prevalence Study was unknown. An adjustment was made to the school base weight to reflect those sampled schools found in nonparticipating districts, which were nonrespondents and whose eligibility status was unknown. Nonresponse to the district recruitment effort was somewhat higher for suburban and rural districts than for urban districts, and for noncertainty districts among the urban districts. (Certainty districts are the very largest districts that were sampled with certainty because of their size.) Hence, three nonresponse adjustment cells were formed (1=urban/certainty, 2=urban/noncertainty, and 3=suburban, rural) and the corresponding adjustment factor for each cell was computed as:

$$F_{\text{District Nonresponse Adj}} = \frac{\sum_{i \in \textit{all sampled schools}} bw_i}{\sum_{i \in \textit{sampled schools in participating districts}.} bw_i}$$

where bw_i is the school base weight for the i-th school. Schools with unknown eligibility were assigned an adjustment factor of 0, because they were represented by the sampled schools with known eligibility. The school base weight was multiplied by its corresponding adjustment factor.

Adjustment for prevention program questionnaire nonresponse. A second nonresponse adjustment factor was calculated for nonresponse to the Prevention Program Questionnaire, so that eligible schools that responded could represent the eligible schools that did not respond, as well as the eligible schools that were not sampled. Nonresponse to the Program Prevention Questionnaire was determined to vary by school size, locale, instructional level, Census region, Census division, and percent minority, which led to the formation of 16

nonresponse adjustment cells. The nonresponse adjustment factor was calculated separately within each cell as:

$$F_{\text{Prevalence Questionnaire Nonresp Adj}} = \frac{\sum_{i \in resp+nonresp.} w_i}{\sum_{i \in resp} w_i}$$

where w_i is the school weight that has been adjusted for schools within nonparticipating districts. A total of 4,700 schools were eligible and responding.

The final school weight can be written as the product of the school base weight, the adjustment factor for schools within districts that did not participate, and the Prevention Program Questionnaire nonresponse adjustment factor:

$$\text{Final School Weight} = \text{School Base Weight} \times F_{\text{District Nonresponse Adj.}} \times F_{\text{Prevention Program Nonresp Adj}}$$

The sum of the final nonresponse-adjusted school weights is 83,391, with a 95 percent confidence interval of [81,680, 85,101]. This is the estimated number of public schools in the U.S. that were eligible for the Prevention Program Questionnaire.

Provider and principal surveys. This section describes the calculation of program weights for use in the analysis of Fidelity Study program data. The purpose of the program weight is to permit the sample of programs to represent the population of research-based programs in schools that were identified as having such programs in the Prevalence Study. The program weight is designed to be used for estimates pertaining to programs (i.e., those where programs are the unit of analysis). It could also be used for some student level estimation in circumstances where the population of interest is students in schools where a specific program among the 19 programs targeted by the provider survey has been offered. (This issue is discussed briefly in section A.6.1.)

The program base weight was first calculated to reflect the program's overall probability of selection, including the sampling of districts, schools, and programs within schools as well as nonresponse at the district and school levels. The program base weight was then adjusted for sampled programs whose eligibility status was unknown, and for nonresponse to the Provider Questionnaire to help control for nonresponse bias. A set of jackknife replicate weights

was also created for each program or school implementation for the calculation of standard errors that reflected the sample design.

Primary target population and sample. The primary target population for the provider survey was the set of 19 research-based curriculum programs in operation during the 2004–05 school year in U.S. public schools. The Prevalence Study was conducted to identify the sampled schools with at least one of these programs. The 1,801 schools that participated in the Prevalence Study and indicated that they offered at least one of these 19 programs received a provider survey questionnaire for each sampled program with the school. In most cases all eligible programs were sampled from a school. However, to reduce respondent burden, a subsample of programs was selected within 134 schools. In all, a total of 2,950 programs in 1,801 schools were included in the sample. A discussion of the sample selection of programs appears within section A.1.3.

Note that the term "program" can be ambiguous. Technically, 19 programs are of interest to the study. Of these 19 programs, 2,950 different programs were sampled.

Program base weights. The base weight for each sampled program implementation among the 2,950 sampled was calculated as the product of the school's final weight (developed for the Prevalence Study) and the inverse of the probability of selection of the program implementation within the school:

Program Base Weight = School Final Weight x [1/probability program was sampled within the school]

In most schools, all eligible programs were taken, so no sampling of programs was needed. In these schools, the second factor in this expression is simply 1.

The program base weights sum to 53,637, the estimated total number of reported programs in the U.S based on the data obtained from the Prevalence study. However, note that many of the sampled programs turned out to be ineligible (i.e., did not represent an actual implementation of the program indicated), so the estimated number of programs implemented "with fidelity" to the underlying basis for the program is substantially less.

Adjustment for nonrespondents whose eligibility status for the study was unknown. The eligibility status for the Provider Questionnaire could not be determined for 357 sampled

programs. An adjustment was made to the program base weight so that programs whose eligibility status was known would account for them. The rate of unknown eligibility varied across a number of variables available for the purpose of nonresponse adjustment: metro status, grade level, minority enrollment level, and Census region. Ten cells were formed for this purpose. A separate adjustment factor was calculated within each cell as:

$$F_{\text{Unknown Elig Adj}} = \frac{\sum_{i \in \textit{all sampled programs.}} pgmbw_i}{\sum_{i \in \textit{eligibility known}} pgmbw_i}$$

where $pgmbw_i$ is the program base weight for the i^{th} sampled program. The program base weight of each sampled program with known eligibility status was multiplied by the adjustment factor associated with the cell of which it was a member.

Provider questionnaire nonresponse adjustment. A second nonresponse adjustment factor was calculated among all sampled programs determined to be eligible for the Provider Questionnaire. This second nonresponse adjustment factor produced a weight that permitted eligible programs in schools that responded to the questionnaire to represent the eligible programs for which no Provider Questionnaire was obtained in addition to the eligible programs that were not sampled. Nonresponse to the provider survey was determined to vary by school minority level, metro status, grade level, and Census region. From these variables eight nonresponse adjustment cells were formed. The nonresponse adjustment factor was calculated separately within each cell as:

$$F_{\text{Provider Questionnaire Nonresp Adj}} = \frac{\sum_{i \in resp+nonresp.} pw_i}{\sum_{i \in resp} pw_i}$$

where pw_i is the program weight that has been adjusted for unknown eligibility. One-thousand programs had a completed Provider Questionnaire.[26] The frequency distribution across the general response disposition categories appears in Table A-9.

[26] Two of the 1,002 programs on which data were collected were incorrectly treated as nonresponses. In addition, 39 programs were excluded from the Fidelity Study analysis because they were on Programs with data on ten or fewer implementations each.

Table A-9. Frequency distribution across the general response disposition categories for the Provider Questionnaire

	Number of sampled programs
Eligible respondents	1,000
Eligible nonrespondents	294
Ineligible	1,299
Unknown eligibility	357
Total	2,950

The final program weight can be expressed as the product of the program base weight, the unknown eligibility adjustment factor, and the Provider Questionnaire nonresponse adjustment factor:

$$\text{Final Program Weight} = \text{Program Base Weight} \times F_{\text{Unknown Elig Adj.}} \times F_{\text{Provider Questionnaire Nonresponse Adj}}$$

The sum of the final nonresponse-adjusted program weights is 26,242, with a 95 percent confidence interval of [24,132, 28,351]. This is the estimated number of programs faithfully implemented within U.S. public schools based on the 1,000 eligible completed Provider Questionnaires.

Variance estimation. The estimation of standard errors must take account of features of the sample design such as stratification, clustering, and weighting. For this purpose, a set of 100 jackknife (JK2) replicate weights was created for each sampled school for variance estimation with software packages designed for the analysis of complex survey data, such as WesVar, SUDAAN, and Stata. Replication methods work by dividing the sample into specially designed replicate subsamples that mirror the design of the full sample, such that the variation among the replicate subsamples can be used to estimate the variance of the full sample estimate. The creation of the replicate weights is described in the next section.

School replicate base weights. Because districts were sampled at the first stage, they were the primary sampling units (PSUs) upon which the creation of "drop-groups" (sometimes also called "varunits" or "varpsus") for the jackknife was based, with the exception of districts sampled with certainty. In certainty districts, the drop-groups were based on groups of schools, because schools were the first stage of probability sampling in these districts.

To create variance strata, the noncertainty districts were sorted in order of sample selection; then consecutive district pairs were numbered from 1 to 100 repeatedly (beginning

with 1 again when 101 was reached). Within each pair, districts were numbered 1, 2, 1, 2, etc. Variance strata were assigned from 1 to 100 based on the numbering of consecutive pairs. New districts were assigned the order of selection of the district they were linked to for weighting purposes. Hence, each variance stratum contains districts from across the entire sample and two varpsus, each of which mirrors the full sample. This is important to ensure enough replicates for stable variance estimation for potential analysis domains (e.g., Census region, metropolitan status, instructional level, and percent minority).

Schools from certainty districts were sorted by the school sampling stratum and school order of selection within stratum. New schools were assigned the stratum of the school they were linked to for weighting purposes. Consecutive pairs of schools were numbered sequentially, beginning where the variance stratum numbering had left off for the noncertainty districts and continuing across stratum boundaries to 100. When 101 was reached, numbering began again with 1. Within each pair, schools were numbered 1, 2, 1, 2, etc.

The result of these sorting and numbering operations was the creation of 100 variance strata for the entire school sample with two varpsus per variance stratum. In each variance stratum, a replicate weight was generated by randomly dropping one varpsu (i.e., assigning the replicate weight for schools in the varpsu to 0) and doubling the school base weights of the schools in the remaining varpsu in the variance stratum. This resulted in 100 JK2 replicate base weights being created for each sampled school.

Nonresponse-adjusted school replicate weights. The nonresponse adjustment factors for both district nonresponse and Prevention Program Questionnaire nonresponse were recalculated for each of the 100 replicates in the same manner as was done for the full-sample school weight, using the replicate base weight and the same adjustment cells, so that the sampling variability in the response rates would be captured in the replicate weights. Each school replicate base weight was multiplied by its corresponding replicate nonresponse adjustment factors to create the final nonresponse-adjusted replicate weights. Hence, standard errors produced from these replicate weights reflect all of the components of sampling error.

Program replicate base weights. A set of 100 JK2 replicate base weights was created for each sampled program by multiplying each final school replicate weight by the program full-sample base weight.

Nonresponse-adjusted program replicate weights. The nonresponse adjustment factors for both unknown eligibility and Provider Questionnaire nonresponse were recalculated for each of the 100 replicates in the same manner as was done for the full-sample program weight, using the replicate base weight and the same adjustment cells, so that the sampling variability in the response rates would be appropriately reflected by the replicate weights. Each program replicate base weight was multiplied by its corresponding replicate nonresponse adjustment factors to create the final nonresponse-adjusted replicate weights. Standard errors produced from these replicate weights hence include all the components of sampling error.

Analysis

Separate analyses were conducted for the Prevalence Study and Fidelity Study. The unit of analysis for the Prevalence Study was the school, whereas it was the program for the Fidelity Study. This section describes the data sources and derived and recoded variables used for each study; it also describes the types of analyses conducted for both studies.

Prevalence study analysis. After a brief discussion of how data from different sources were used in analyses, this section covers derived and recoded variables.

Use of data from different sources. The data used in the Prevalence Study were from three sources: (a) Prevention Program Questionnaire, (b) District Questionnaire,[27] and (c) the U.S. Department of Education's 2003–04 nonfiscal Common Core of Data (CCD).

The school is the unit of analysis for this study, with the data weighted to make estimates at the school level. However, information on programs and districts also was used in analyses. The information on programs (e.g., number of programs supported by SDFSCA Program funding) comes from aggregating weighted school-level responses regarding the number and type of prevention programs. For example, to obtain estimates on the number of prevention programs implemented nationwide, a count was obtained for each school by adding the number of different prevention programs the respondent reported using in the school. If a respondent indicated that 10 different programs were used in the school and the school weight was equal to 14 (i.e., this particular sampled school represented 14 schools in the nation), then

[27] The following District Questionnaire items were used as single-item indicators in the analysis: Q13, Q21A-Q21E, Q23A-Q23E, Q26A, and Q26B.

140 programs were counted for the school. This procedure was carried out for each school, and the weighted number of programs was summed across all schools to obtain the aggregated weighted total.

Information obtained from the district survey was merged into the corresponding school records. However, not all of the schools had district information, due to district survey nonresponse. In cases where the district data could be merged onto the school data, statements were made about schools in districts with certain characteristics. For example, analysts examined the relationship between school use of research-based programs and district coordinator time devoted to SDFSCA Program funded activities. By conducting the analyses with both types of data, analysts were able to conclude that schools in districts in which the district coordinator spent little to no (0 to 5 percent) time on SDFSCA Program activities were less likely to use one or more research-based programs compared to schools in districts in which the coordinator spent more than 5 percent time on SDFSCA Program activities.

Derived and recoded variables. The derived and recoded variables used in the study analyses are described in this section. Definitions of each variable are provided along with additional variable construction information, when appropriate.

School enrollment (all schools)—This measure is the total number of students enrolled in school based on data from the 2003–2004 CCD. The variable was collapsed into the following three categories:

- *Less than 300 students*
- *300 to 999 students*
- *1,000 or more students*

Elementary school enrollment—Total number of students enrolled in elementary schools based on data from the 2003–04 CCD. The variable was collapsed into the following three categories:

- *340 or less students*
- *341 to 530 students*
- *More than 530 students*

Middle school enrollment—Total number of students enrolled in middle schools based on data from the 2003–04 CCD. The variable was collapsed into the following three categories:

- *Less than 440 students*
- *441 to 760 students*
- *More than 760 students*

High school enrollment—Total number of students enrolled in high schools based on data from the 2003–04 CCD. The variable was collapsed into the following three categories:

- *325 or less students*
- *326 to 960 students*
- *More than 960 students*

Instructional level—Using the 2003–04 CCD data, this variable was created from the lowest and highest grades for which students were reported in a school. The NCES definition of instructional level was applied to generate the following four levels:

- *Elementary schools: Low grade of prekindergarten through grade 3, high grade of up to grade 8*
- *Middle schools: Low grade of 4 to 7, high grade ranging from grades 4 to 9*
- *High schools: Low grade of 7 to 12, high grade must extend through grade 12*
- *Other schools: All other grade configurations, including schools that were completely ungraded*

Percent minority enrollment—This variable indicates the percentage of students enrolled in a school whose race or ethnicity were classified as black or Hispanic, based on data from the 2003–04 CCD. The variable was collapsed into the following three categories:

- *Less than 11 percent minority enrollment*
- *11 to 60 percent minority enrollment*
- *More than 60 percent minority enrollment*

School urbanicity—This variable reflected the type of community in which the school was located as defined by the 2003–04 CCD LOCALE03 variable. The original eight category variable was recoded into three categories. The new categories of urban, suburban, and rural have the following definitions:

- *Urban—Large or mid-size principal city of a Metropolitan Core Based Statistical Area (MCBSA)*
- *Suburban—Urban fringe of a large or mid-size city*
- *Rural—Large town (population greater than or equal to 25,000 and located outside a MCBSA), small town (population less than 25,000 but greater than or equal to 2,500 located outside a MCBSA), or any rural area as defined by the Census Bureau inside or outside an MCBSA*

School percent eligible for free and reduced-price school lunches—This variable was created from the following two CCD variables: (a) TOTFRL03, a count of the total of free lunch eligible and reduced-price lunch eligible students in the school and (b) MEMBER03 a count of the total school membership. The percentage of students eligible for free and reduced-price school lunches was computed as TOTFRL03 divided by MEMBER03 then multiplied by 100. The variable was collapsed into the following three categories:

- *25 percent or less of students*
- *26 percent to 55 percent of students*
- *More than 55 percent of students*

District enrollment—Information on district enrollment was obtained from question 33 on the District Questionnaire, which asked respondents, "As of October 2, 2004, how many students were enrolled in your district?" The responses were collapsed into the following three categories:

- *Less than 3,000 students*
- *3,000 to 14,999 students*
- *15,000 or more students*

District percent eligible for free and reduced-price school lunches—Data obtained for this variable came from question 34 of the District Questionnaire, which asked, "As

of October 1, 2004, what percentage of students in your district were eligible for free or reduced-price lunch benefits?" The responses were collapsed into the following three categories:

- *30 percent or less of students*
- *31 percent to 50 percent of students*
- *More than 50 percent of students*

Number of prevention programs—The Prevention Program Questionnaire asked respondents to identify whether their school had used each of 103 programs during the 2004–05 school year. The survey also contained 39 "other specify" (open-ended) items in which the respondent could write in a program not listed. When a respondent mistakenly wrote in a program already listed, the response was recoded as the listed program. All respondents started with a zero and received a 1 for each program used in the school. Although a respondent could indicate that his or her school implemented as many as 142 programs, the number of programs reported by respondents ranged from 0 to 72. For some analyses, the number of prevention programs was collapsed into the following six categories:

- *0 programs*
- *1 to 5 programs*
- *6 to 10 programs*
- *11 to 15 programs*
- *16 to 20 programs*
- *More than 20 programs*

Program type—Each program listed in the Prevention Program Questionnaire fell into one of 13 program types. The number of programs under each type is simply a count of how many programs within each type a respondent indicated his or her school used during the 2004–05 school year. These counts by program type were summed to calculate the total number of prevention programs across all program types.

Number of research-based programs by program type—The 21 Research-based Programs identified for the Prevalence Study can be grouped into three of the 13 program types. The number of research-based programs by program type consists of three measures that are

simply how many research-based programs within each type (i.e., prevention curriculum, school climate and programs for families) a respondent indicated his or her school used during the 2004–05 school year. For the total number of research-based programs, these counts for all three program types were summed.

Received funding from SDFSCA Program (school-level)—The Prevention Program Questionnaire asked whether a program was used in the school and whether the "school received any funds from the federal SDFSCA Program to support the program." Response options were, "yes," "no," and "don't know." A count was conducted for each school on the number of programs receiving SDFSCA Program funds and the number of programs not receiving SDFSCA Program funds (i.e., number of programs for which the respondent answered "no"); these two counts were summed for the total number of programs respondents indicated were funded or not funded by the SDFSA Program. In estimates of the percentage of schools using programs funded by the SDFSCA Program, the denominator included only the respondents who knew whether a given program was funded by the SDFSCA Program.

Number of programs by whether research-based and funded by SDFSCA Program—In order to determine the number of prevention programs by whether or not they were research-based and received SDFSCA Program funding, counts were conducted to create the following classifications:

- *Number of research-based programs that received SDFSCA Program funding*
- *Number of research-based programs that did not receive SDFSCA Program funding*
- *Number of non research-based programs that received SDFSCA Program funding*
- *Number of non research-based programs that did not receive SDFSCA Program funding*
- *Total number research-based programs reported by respondents who knew whether or not the programs received SDFSCA Program funding*
- *Total number of non-research-based programs reported by respondents who knew whether or not the programs received SDFSCA Program funding*

Most frequently reported "named" prevention programs—"Named" prevention programs refer to: (a) programs listed on the Prevention Program Questionnaire, (b) Helping America's Youth[28] programs that were written in the "other specify" fields, and (c) the DARE program. In total, the list of "named" programs includes 103 programs. Of these 103 programs, the most frequently reported "named" programs are highlighted in Table 3 of the report.

Total number of "named" programs—This measure is a count of programs implemented during the 2004–05 school year that are listed on the Prevention Program Questionnaire, in addition to any HAY programs and DARE programs written in the "other specify" fields of the questionnaire. Any programs indicated in two or more locations on the Prevention Program Questionnaire were counted only once. [29]

Number of research-based programs—A program is considered "research-based" if it is one of 21 programs identified as effective through an extensive screening and research review process (described in Chapter 2). A count was conducted for each school on the number of programs that were research-based and the number of programs that were not research-based. Each school could receive a count ranging from 0 to 21 that reflected the number of research-based programs used in the 2004–05 school year. However, the actual count of research-based programs ranges from 0 to 10. The most frequently reported research-based programs are highlighted in Table 4 of the report. For some analyses, the number of research-based programs was collapsed into the following four categories:

- *no programs*
- *1 program*

[28] Helping America's Youth (HAY) was a nationwide effort led by Mrs. Laura Bush and supported by the White House, U.S. Department of Justice, and several other federal agencies. Its goals were to raise awareness about the challenges facing youths and to motivate caring adults to connect with these youths. HAY promoted the use of programs that are said to "have been evaluated and found to successfully deal with risky behaviors." The following HAY programs were not included on the Program Questionnaire but were included in the count of "named" programs: Academic Tutoring and Social Skills Training, American Indian LifeSkills Development, Families in Action, Head On, Not on Tobacco (NOT), Project EX, SMART Talk (Students Managing Anger and Resolution Together), Spit Tobacco Intervention for High School Athletes, Students Training Through Urban Strategies (STATUS), Success in Stages, and Too Good for Violence.

[29] These programs included the following: Classroom-Centered (CC) and Family-School Partnership (FSP) Intervention, Consistency Management and Cooperative Discipline (CMCD), East Texas Experimental Learning Center, Families in Action, Incredible Years, Leadership and Resiliency (LRP), PeaceBuilders, Proactive Classroom Management, Project Northland, Project PATHE (Positive [Promoting] Action Through Holistic Education), Project SUCCESS, School Safety Program, Schools and Families Educating Children (SAFE Children), Seattle Social Development Project/Skills, Opportunities, and Recognition (SOAR), Spit Tobacco Intervention for High School Athletes, and Students Training Through Urban Strategies (STATUS).

- *2 programs*
- *3 or more programs*

For additional analyses, the number of research-based programs was collapsed into a dichotomous variable where "one or more research-based programs" was coded as 1 and "no research-based programs" was coded as "0."

Percent of district coordinator time devoted to SDFSCA activities—Question 29a on the District Questionnaire asked respondents to indicate approximately what percentage of their time was devoted to activities receiving SDFSCA Program funding during the 2004–05 school year. Responses ranged from 0 to 100 percent. This variable was collapsed into the following categories:

- *0 to 5 percent*
- *6 percent to 25 percent*
- *More than 25 percent*

Time district coordinator involved with prevention (years)—This variable is measured using question 30 on the District Questionnaire. It asked, "Counting the 2004–2005 school year, for how many years have you worked on district-level prevention activities?" Responses ranged from 0 to 37 years. For analysis, the time a district coordinator was involved with prevention activities was collapsed into the following three categories:

- *0 to 4 years*
- *5 to 9 years*
- *10 or more years*

Length of time since district coordinator attended prevention workshop/conference—Question 32 of the District Questionnaire asked respondents, "In what year did you last receive training or attend a conference/workshop focused specifically on ATOD use or school crime prevention?" This open-ended question also allowed respondents to check a box indicating they had never received training or attended a conference or workshop focused on

those topics. Responses ranged from 1982 to 2006. This variable was collapsed into the following four categories:

- *Never*
- *1 year ago or less*
- *2 to 5 years ago*
- *More than 5 years ago*

Level of parent and community involvement in district prevention planning—Parent and community involvement was measured using responses from the following District Questionnaire items that assessed how involved parents and community members were in: 1d) analyzing data on problem behavior, 1e) reviewing prevention goals and objectives, and 1f) selecting prevention programs. Response options were coded very much=4, somewhat=3, not very much=2, and not at all=1. The sum of the three items, taking into account missing data, resulted in respondent scores ranging from 1 to 12. These scores were then recoded to reflect the level of parent community involvement where 1 to 3 = none, 4 to 6 = a little, 7 to 9 = some, and 10 to 12 = a lot.

District considered student needs in prevention needs assessment—This variable was measured using question 12l on the District Questionnaire. It asked respondents to indicate how much district data on student needs was a factor in adding or dropping ATOD use and school crime prevention programs. Response choices were, "very much," "somewhat," "not very much," and "not at all." For analysis, these response options were recoded to create a dichotomous variable with a yes or no response format.

Fidelity study analysis. The analyses for the Fidelity Study focused on a subset of the 19 Research-Based Programs on which data were collected. Although original plans called for separately analyzing data from the three program types on which providers were surveyed (prevention curriculum, instruction, or training programs; programs to change or maintain the culture or climate of the school, alter or maintain expectations for student behavior, or secure commitment to norms; and programs or services for families or family members), the sample sizes and responses from only the prevention curriculum, instruction, or training program type made that the only feasible program to assess. Hence, the analyses excluded programs intended to change or maintain the culture or climate of the school, alter or maintain expectations for student behavior, or secure commitment to norms; and programs or services for families or

family members. Moreover, only 10 of the prevention curriculum, instruction, or training Programs were included in the analyses. These Programs were selected because data were available on over 10 eligible implementations of them. They are: (a) Aggression Replacement Training, (b) Alcohol Misuse Prevention, (c) All Stars, (d) Anger Coping, (e) Know Your Body, (g) Life Skills Training, (h) Positive Action, (i) Project Alert, (j) Promoting Alternative Thinking Strategies, and (k) Second Step.

After a brief discussion of how data from different sources were used in analyses, this section also describes the development of the study's fidelity measures and standards and of derived and recoded variables

Use of data from different sources. The data used in the Fidelity Study were from six sources: (a) Prevention Program Questionnaire, (b) District Questionnaire,[30] (c) Provider Questionnaire, (d) Principal Questionnaire, (e) fidelity standards developed based on a review of program materials, and (f) the U.S. Department of Education's 2003–04 nonfiscal Common Core of Data (CCD).

The program is the unit of analysis for this study, with the data weighted to make estimates at the Program level. Information on district and school characteristics also was used in analyses. Hence, statements were made about programs in schools and districts with certain characteristics. For example, analysts examined the relationship between fidelity on number of lessons provided and district coordinator time devoted to SDFSCA Program funded activities. By conducting the analyses with both types of data, analysts were able to examine whether programs in districts in which the district coordinator spent little to no (0 to 5 percent) time on SDFSCA Program activities were more or less likely to meet fidelity standards on number of lessons provided than programs in districts in which the coordinator spent more than 5 percent time on SDFSCA Program activities.

One should note that, while the analyses in this report are all at the program level, analyses that are enrollment based may also be of interest. For example, the proportion of all students who are in schools offering a particular program that is being implemented with fidelity may differ from the proportion of programs that is being implemented with fidelity. This could arise, for example, if programs being implemented with fidelity are more highly concentrated in rural schools which tend to be smaller than their more urban counterparts.

[30] The following District Questionnaire items were used as single-item indicators in the analysis: Q21D and Q21E.

One should also note that any analyses incorporating enrollment that use the Fidelity Study data can be undertaken only for the individual program types rather than for all programs. This is so because a number of schools offer multiple programs and, to limit respondent burden, subsampling among programs within a school was occasionally necessary. One cannot consider enrollment at the program level across all programs because some schools offer multiple programs, and this would result in double counting. One cannot consider enrollment at the school level (i.e., schools with at least one eligible program, so that a single enrollment figure could be used) because schools with programs that were subsampled may have an eligible program that was not among those that happened to be subsampled.

Development of fidelity measures and standards. To assess the fidelity of implementation, measures were developed against which program provider responses on program implementation could be compared. For the purposes of this study, two types of measures were developed: program-specific measures and generic measures of implementation fidelity. For both types of measures, the study's researchers established thresholds or cut points, in consultation with the Office of Safe and Drug-Free Schools, for deciding whether the implementation of a given program passed or failed on a specific aspect of implementation. These thresholds or cut points are known as fidelity standards.

Development of the program-specific fidelity standards relied primarily on the instructional materials created by the Program developers. A full set of materials was ordered from each program publisher, including instructional guides, student workbooks, manipulatives, and audiovisual materials. Implementation fidelity information was gathered primarily from instructional guides (e.g., teachers' guides or lesson plans), but some fidelity information was found through searching ancillary sources, such as the Program developer's website, online directories of prevention programs (e.g., SAMHSA Model Programs), and published research literature. Attempts were made to obtain materials that were published prior to the 2004–05 school year, but in some cases only newer material was available.

Staff reviewed the program materials for implementation fidelity information. The tool used for capturing standards of fidelity was a list of items from the Provider Questionnaire. Although the program reviews were conducted independently by the two reviewers, regular meetings were held concurrently with early reviews to check inter-rater consistency and discuss the application of the fidelity constructs to the program materials.

The reviewers first searched a Program's materials for information relevant to each measure and recorded the appropriate fidelity standard (or the absence of a standard). To take into account variation in program fidelity standards across levels of a Program, reviews were performed for each separately published Program component. For example, Programs with individual curricula for different grades, or those with a booster module, required separate reviews for each program unit. If a component of a Program was identified by the program materials as especially important to successful implementation, the reviewers flagged the item as a central aspect of the Program. How these central items are handled in analysis is described in Section A.6.3.

Review of developer and other materials for some specific standards was unfruitful, because the materials were silent on the standard. If this occurred for only one or a few Programs, the standards were used for the Programs on which information was available. However, if information on standards was unavailable for most of the Programs, those standards were dropped from consideration in developing measures of fidelity. The aspects of implementation for which standards could be developed are listed in Table A-10, along with the Provider Questionnaire items used to capture information on the standards for specific programs.

In addition to program-specific measures of implementation fidelity, measures of generic implementation fidelity were developed. These measures, which cut across Programs, were based on a review of meta-analyses and expert judgment (see Section A.2.1.2).

Table A-10. Standards for measuring fidelity of implementation

Standards	Provider Questionnaire item number
Frequency of student participation	Q12
Number of lessons	Q13
Topics covered	Q14
Targeting on level of risk	Q32

Derived and recoded variables. The derived, recoded and other variables used in the study analyses are described in this section. Definitions of each variable are provided along with additional variable construction information, where appropriate.

Fidelity standard for required topics covered—This fidelity measure is based on criteria found in a review of program implementation manuals and other related materials regarding topics that instructors are expected to cover during program implementation. The values for this measure resulted from a count of specific topics found in implementation

materials that correspond to the response options provided to respondents in Q14 of the Provider Questionnaire. This measure represents 70 percent of the manual standard for topics covered. See also description of *topics covered*.

Topics covered—This measure was intended to assess whether implementation of a given program followed the prescription of the program developer for specific topics that should be covered by program providers. Question 14 on the Provider Questionnaire asked respondents to indicate whether specific topics were actually taught by instructors of the program of interest during the 2004–05 school year. Responses to this question were compared to the Fidelity Study standard for each program, which was based on information found in the program manual(s) or other related implementation materials. If a respondent indicated that program providers covered at least at least a minimum number of specific topics (70 percent of the prescribed number of specific topics), the implementation "passed" on *topics covered*.

Fidelity standard for number of lessons provided—This fidelity measure is based on criteria found in a review of program implementation manuals and other related materials regarding the number of required lessons. This measure represents 70 percent of the manual standard for number of lessons. See also description of *number of lessons*.

Number of lessons—This measure was intended to assess whether implementation of a given program followed the prescription of the program developer for the number of lessons that should be provided to participating students. Question 13 on the Provider Questionnaire asked respondents how many lessons most participating students completed. Responses to this question were compared to the Fidelity Study standard for each program, which was based on information found in the program manual(s) or other related implementation materials. If a respondent indicated that participating students completed at least a minimum number required of lessons (70 percent of the prescribed number of lessons), the implementation "passed" on *number of lessons*.

Fidelity standard for frequency of student participation—This fidelity measure is based on criteria found in a review of program implementation manuals and other related materials regarding frequency of student participation. The values for this measure correspond to the response options provided to respondents in Q12 of the Provider Questionnaire. See also description of *frequency of student participation*.

Frequency of student participation—This measure was intended to assess whether implementation of a given program followed the prescription of the program developer for how often students should participate. Question 12 on the Provider Questionnaire asked respondents how often most students participated in the program of interest during the 2004–05 school year. Respondents were asked to select one from a list of eight response options: (a) More than once a day, (b) Once a day, (c) 2 to 6 times a week, (d) Once a week, (e) 2 or 3 times a month, (f) Once a month, (g) Less than once a month, and (h) Once or twice during a school year. Responses to question 12 were compared to the Fidelity Study standard for each program, which was based on information found in the implementation manual for the program. If a respondent indicated that most students participated in the program of interest with frequency greater than or equal to the Fidelity Study standard, the implementation "passed" on frequency of student participation.

Fidelity standard for level of risk targeted (1 and 2)—These fidelity measures are based on criteria found in program implementation manuals and other related materials or as classified by a federal agency (e.g., SAMHSA) as appropriate levels of risk to be targeted. These levels of risk correspond to the Institute of Medicine's classifications of universal, selective, and indicated levels of risk. The first standard is the standard for programs that target one level of risk and the second standard is for programs that target more than one level of risk. See also description of *level of risk targeted*.

Level of risk targeted—This measure was intended to assess whether implementation of a given program followed the prescription of the program developer for the level of risk for the targeted population (i.e., universal, selective, or indicated). Question 32 on the Provider Questionnaire asked respondents to indicate which of three populations (varying on level of risk) best described students who participated in the program of interest during the 2004–05 school year. Responses to this question were compared to the Fidelity Study standard for each program, which was based on information found in the program manual(s) or other related implementation materials. If a respondent indicated a population served that was consistent with the Fidelity Study standard, the implementation "passed" on *level of risk targeted*.

Generic fidelity standard for use of instructional strategies—This fidelity measure is the count of "yes" responses to items 15g, 15h, 15j, 15l, 15p, and 15q on the Provider Questionnaire. Question 15 on the Provider Questionnaire asked respondents to indicate whether or not specific instructional strategies were used by instructors of the program during the 2004–05 school year. Six of these 21 items (i.e., 15g, 15h, 15j, 15l, 15p, and 15q) included in this

question are descriptions of instructional strategies considered to be best practices for method of delivery (Gottfredson et al., 2000; and review of meta-analyses conducted for this study). This measure sets the minimum requirement for quality program implementation at 70 percent (i.e., at least 4 out of 6 items).

Instructional strategies—This measure was intended to assess whether implementation of a given program incorporated instructional strategies that are considered best practices for method of delivery. Question 15 on the Provider Questionnaire asked respondents to indicate whether or not specific instructional strategies were used by instructors of the program of interest during the 2004–05 school year. Six of the 21 items included in this question (i.e., behavioral modeling, role-playing, practice of new skills, use of cues to encourage certain behaviors, behavioral management or behavioral modification techniques, and peer teachers or leaders) are considered best practices for method of delivery (Gottfredson et al., 2000). Affirmative responses to question items that were consistent with these best practice instructional strategies (i.e., 15g, 15h, 15j, 15l, 15p and 15q) were added and the sum of employed strategies was compared to the Fidelity Study standard for each program. If a respondent indicated that program providers employed at least at least a minimum number of instructional strategies (i.e., 4 out of 6), the implementation "passed" on *instructional strategies***.**

Generic fidelity standard for rewards, recognition, and student mastery—This fidelity measure corresponds to the affirmative responses to items Q19b, Q19c, Q20b, Q20c, Q20d, Q21a, and Q21b and the negative response to Q20e on the Provider Questionnaire. These items represent methods of rewards, recognition and student mastery considered best practices for method of delivery (Gottfredson et al., 2000). This measure sets the minimum requirement for quality program implementation at 70 percent (i.e., at least 6 out of 8 items).

Rewards, recognition, and mastery assessment—This measure was intended to assess whether implementation of a given program incorporated specific reward, recognition, and assessment of student mastery methods that are considered best practices for method of delivery (Gottfredson et al., 2000). The items that comprise this measure include eight items from three questions on the Provider Questionnaire: Two items from question 19 (i.e., application of rewards for individual and group achievements), four items from question 20 (i.e., student recognition for effort, improvement, successful competition and performance), and two items from question 21 (i.e., assessment of student mastery and re-teaching material not mastered). Responses to questions 19, 20, and 21 that are consistent with these "best practices" methods were added and the sum was compared to the Fidelity Study standard for each program.

If a respondent indicated that program providers employed at least a minimum number of best practices methods (i.e., 6 of 8), the implementation "passed" on *rewards, recognition, and mastery assessment*.

Number of standards passed—This measure was derived by creating a composite score (ranging from 0–6) that summed "pass" or affirmative values (i.e., "1") for the six implementation fidelity measures. These measures included the following:

- *Frequency of student participation*
- *Level of risk targeted*
- *Number of lessons*
- *Instructional strategies*
- *Rewards, recognition and mastery assessment*
- *Topics covered*

Overall implementation fidelity—This dichotomous measure was intended to establish a minimum threshold for overall implementation fidelity. Programs that "passed" on at least the four program-specific standards (i.e., *frequency of student participation, level of risk targeted, number of lessons, topics covered*) were assigned the status of "adequate" *overall implementation fidelity*; Programs that did not pass on at least the four program-specific standards were assigned the status of "inadequate" *overall implementation fidelity.*

Instructional level—This variable was also used in the Prevalence Study. See description provided in Section A.6.1.2.

Received funding from SDFSCA Program (program-level)—This measure was derived from 10 different SDFSA Program funding variables from the prevention program survey. Respondents were asked in questions 1-13 (item d) of the prevention program survey to indicate whether or not specific prevention programs received any support from federal SDFSCA Program funds during the 2004–05 school year. Question 1(d) asked about SDFS funding received for specific "Behavioral Programming or Behavior Modification Programs" that included Know Your Body. Question 3 (item d) asked about SDFS funding received for specific "Prevention Curriculum, Instruction, or Training Programs" that included Adolescent Alcohol Prevention Trial (AAPT)/All Stars, Aggression Replacement Training, Alcohol Misuse

Prevention, Anger Coping Program, Life Skills Training, Positive Action, Project ALERT, Promoting Alternative Thinking Strategies (PATHS), and Second Step. Respondents were instructed to indicate their answers using a "yes," "no," "don't know" response format.

School enrollment (all schools)—This variable was also used in the prevalence Study. See description provided in Section A.6.1.2.

School percent eligible for free and reduced-price school lunches—This variable was also used in the Prevalence Study. See description provided in Section A.6.1.2.

School urbanicity—This variable was also used in the Prevalence Study. See description provided in Section A.6.1.2.

District enrollment—Information on district enrollment was obtained from question 33 on the District Questionnaire which asked respondents, "As of October 2, 2004, how many students were enrolled in your district?" The responses were collapsed into the following three categories:

- *Less than 3,000 students*
- *3,000 to 14,999 students*
- *15,000 or more students*

District percent eligible for free and reduced-price school lunches—This variable was also used in the Prevalence Study. See description provided in Section A.6.1.2.

Percent of coordinator time devoted to SDFSCA Program activities—This variable was also used in the Prevalence Study. See description provided in Section A.6.1.2.

Length of time since district coordinator attended a prevention workshop or conference—This variable was also used in the Prevalence Study. See description provided in Section A.6.1.2.

Extent of district monitoring—Respondents were asked (in question 16 on the District Questionnaire) to indicate whether or not the district monitored certain aspects of program implementation for programs with SDFSCA Program funding and programs without

such funding. Using a "yes" "no" response format, this question included four items that were worded as follows: "(a) Extent to which planned program "dosage" or exposure was achieved, (b) Extent to which the content of prevention activities was adapted, (c) Number or characteristics of students served, and (d) Progress toward the achievement of outcome goals or objectives for prevention activities." Responses for programs with SDFSCA Program funding and those without such funding were combined into a single score using the following coding scheme: If a respondent indicated "yes" on any of the four aspects of monitoring for programs with SDFSCA Program funding or those without such funding, this aspect (e.g., question 16a, extent to which planned program dosage or exposure was achieved) was coded as "1." Coded "yes" responses for items 16a through 16d were summed to create a composite score ranging from 0-4.

Programs evaluated by school or district—This measure was created from questions 19a and 19b on the District Questionnaire. Question 19a asked respondents to indicate who had primary responsibility for evaluating prevention programs receiving SDFSCA Program funding during the 2004–05 school year. Question 19b on the District Questionnaire asked respondents who had primary responsibility during the 2004–05 school year for evaluating prevention programs that were not funded by the SDFSCA Program. Response options included: *district staff*=1, *school staff*=2, *equally shared district and school responsibility*=3, and *not done during the 2004–2005 school year*=4. EVALPROG was coded as "yes=1" if a respondent provided an affirmative response (i.e., 1, 2, or 3) to either questions 19a or 19b. EVALPROG was coded as "no=2" if a respondent indicated a negative response (i.e., 4) to either questions 19a or 19b.

Quality of initial training—Question 40 on the Provider Questionnaire asked respondents to describe the initial training for the program. Using a "yes," "no" response format, the question items were worded as follows:

a. *The presentation was clear and organized*
b. *Principles to be followed were presented*
c. *Principles were illustrated with examples*
d. *Participants practiced applying the principles*
e. *Participants received feedback on their performance in applying the principles*

f. Participants' questions and concerns about possible obstacles in applying the principles were addressed

g. Trainers provided assistance to participants in solving problems or implementing new practices following the training

h. As part of training, participants made specific plans (or set goals or standards) for the program

i. Participants were asked to make public commitments to use specific new methods as part of the training

j. Participants were provided with manuals, notebooks, workbooks, or recording forms to take back to the school or classroom to assist in putting the program in place

This question also allowed respondents to check a box indicating that they did not attend the initial training. For respondents who did attend the initial training, responses were coded yes=1 and no=2. Coded "yes" responses for items 40a through 40j were summed to create a composite score ranging from 0-10. Composite scores (ranging from 0-10) were recoded into three categories so that scores ranged from 0 to 3 = 1 (Low), 4 to 7 = 2 (Medium), and 8 to 10 = 3 (High).

Amount of program specific training—This measure was derived from questions 38, 41, and 42 on the Provider Questionnaire. Responses to each of these three questions were recoded into one of two categories and affirmative responses were coded "1." Affirmative responses were then summed to create composite scores ranging from 0 to 4. These scores were then collapsed into three categories of low (0 to 1), medium (2), and high (3). Questions 38, 41, and 42 are described as follows.

> Question 38 on the Provider Questionnaire asked respondents, "How much training on this program was completed by most instructors the first time it was conducted at the school?" This question allowed respondents to check a box indicating that there was no initial training. Respondents were asked to select one of the following responses regarding the amount of initial training that was completed by most instructors of the program:
>
> *a. 4 days or more*
>
> *b. 2 – 3 days*
>
> *c. 1 full day*

d. A half day

e. Short demonstration or orientation only

Response options were recoded so that *4 days or more*, *2 – 3 days*, *1 full day*, and *A half day* = 1, and *Short demonstration or orientation only* = 0. If a respondent checked the box indicating that no initial training occurred, this response was recoded as "0."

- Question 41 on the Provider Questionnaire asked respondents how many times during the 2004–05 school year that formal follow-up training was completed by most instructors providing the program. Respondents were asked to select one of the following response options:

  a. None

  b. Once

  c. Twice

  d. Three or more times.

  Responses were recoded so that *None* = 0, AND *Once, Twice*, and *Three or more times* = 1.

- Question 42 asked respondents to indicate whether or not ongoing coaching or mentoring from specific sources was available for instructors conducting the program during the 2004–05 school year. Using a "yes" "no" response format, this question included five items that were worded as follows:

  a. Staff from this school

  b. Staff from another school in this district

  c. School district staff

  d. Trainer(s) from the program developer/publisher

  e. Other source (Please specify)

  The items for question 42 were collapsed to create a single measure for "ongoing coaching or mentoring." If a respondent indicated "yes" on any of the five items listed in question 42, ongoing coaching or mentoring were coded as "1." If a respondent did not respond to all five items, or indicated "no" to all of the five sources (or if "no" responses were combined with responses not ascertained), ongoing coaching or mentoring was coded as "0."

Analyses

This study used univariate and bivariate analyses to develop national estimates in answering the main research questions. All analyses were performed using WesVar, a statistical software application appropriate for use with complex sample designs and weighted data. The univariate analyses included frequencies (counts and percentages), means, and 95 percent confidence intervals surrounding the estimates.

Bivariate analyses consisted of crosstabulations in which program estimates were made in conjunction with a second variable. For example, the weighted number and percentage of prevention programs passing on a given measure of fidelity is examined at each of four instructional levels. *Chi-square* tests of independence were conducted on categorical bivariate program-level estimates using the *Rao-Scott (RS3) chi-square* approximation. The *RS3 chi-square* is a modified statistic that reflects the complex sample design used. This modified *chi-square* statistic relies on adjusting *Pearson's chi-square* by using an estimated "design effect" (Westat, 2000).

REFERENCES

Bangert-Drowns, R.L. (1988). The effects of school-based substance abuse education—A meta-analysis. *Journal of Drug Education,* 18(3), 243–264.

Bruvold, W.H. (1993). A meta-analysis of adolescent smoking prevention programs. *American Journal of Public Health,* 83(6), 872–880.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Flay, B. Biglan, A., Boruch, R., Castro, F., Gottfredson, D., Kellam, S., Moscicki, E., Schinke, S., Valentine, J., and Ji, P. (2005). Standards of Evidence: Criteria for efficacy, effectiveness, and dissemination. *Prevention Science,* 6(3), 151–175.

Gottfredson, G.D., Gottfredson, D.C., Czeh, E.R., Cantor, D., Crosse, S.B., and Hantman, I. (2000). *National Study of Delinquency Prevention in Schools: Final Report.* Ellicott City, MD: Gottfredson Associates.

Gottfredson, D.C., and Wilson, D.B. (2003). Characteristics of effective school-based substance abuse prevention. *Prevention Science,* 4(1), 27–38.

Mihalic, S. (2003). *Matrix of programs as identified by various federal and private agencies.* Retrieved from http://www.colorado.edu/cspv/blueprints/matrix/overview.html on Oct. 15, 2003.

Ringwalt, C.L., Ennett, S., Vincus, A., Thorne, J., Rohrbach, L.A., and Simons-Rudolph, A. (2002). The prevalence of effective substance use prevention curricula in U.S. middle schools. *Prevention Science* 3(4), 257–265.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper and L.V. Hedges (eds.), *The Handbook of Research Synthesis* (pp. 231-244). New York: Russell Sage Foundation.

Tobler, N.S.R., Roona, M., Ochshorn, P., Marshall, D.G., Streke, A.V., and Stackpole, K.M. (2000). School-based adolescent drug prevention programs: 1998 Meta-analysis. *Journal of Primary Prevention,* 20(4), 275–336.

Westat (2000). *WesVar 4.0 User's Guide*. Westat: Rockville, MD.

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.

APPENDIX B
SUMMARY REPORT ON THE RESEARCH SYNTHESIS

In this brief report, we describe our approach to the research synthesis and summarize the results from it. The identification of the research base on prevention program effectiveness occurred between May 2004 and December 2005. First, we provide background information on the synthesis.

Background

The evaluation questions for the Study of the Implementation of Research-Based Programs and Practices to Prevent Youth Substance Abuse and School Crime encompassed research-based programs and practices. Specifically, the study sought to answer the following two research questions:

- What proportion of drug and/or violence prevention programs nationally, and in the SDFSCA Program, are implementing **research-based** drug and/or violence prevention programs and practices that scientific evidence has shown produce positive outcomes?

- To what extent nationally, and in the SDFSCA Program, are drug and/or violence prevention programs that are implementing **research-based** programs and practices doing so with fidelity to the research on which they are based?

Identification of Research-based Programs

An initial review of literature indicated that identifying specific research-based programs and general practices required separate approaches. Developing a list of specific Programs entailed identifying and reviewing over 2,000 individual study reports on programs that were judged to be effective by external sources. Developing a list of general practices involved review of a limited number of meta-analyses that provided quantitative results, across many studies, on the effectiveness of practices and general program types; it also drew on the list of research-based practices identified by Gottfredson et al. (2000).

Developing a valid and useful list of Programs for the study entailed compiling and screening existing lists of Research-based prevention Programs, assessing the quality of evidence on the Programs that pass the screens, and making judgments on whether high quality evidence on a given Program indicated a pattern of program-related effects. In the subsections that follow, this review process is discussed in terms of completing six steps (see Figure B-1).

Compiling and screening lists of program. The Prevalence Study decided against using existing lists of programs with stringent inclusion criteria for several reasons. First, those inclusion criteria are often vague, which made assessing their stringency difficult. Second, the U.S. Department of Education was concerned that even lists that used less stringent criteria could include both highly effective and less effective programs; hence, dropping such lists could lead to the unwarranted elimination of programs that should be included on the list for the study.

To be comprehensive, the specification of Research-Based Programs began with a master list of programs developed by Mihalic (2003). This master list aggregates 12 existing lists of programs intended to prevent problem behavior. (For the 12 lists of programs, see Exhibit B-1.) The aggregated master list contained 306 programs; after reviewing the individual lists for more recent information (through 2004), 11 programs were added, for a total of 317 programs.

Step one of the review process entailed identifying the programs on the aggregated master list that were most closely aligned with the mission of the Safe and Drug-Free Schools and Communities Program. Two research staff independently screened each of the programs for whether they were (a) entirely school-based or had separable components that were school-based, (b) focused on the prevention of youth substance abuse or school crime (including violence and aggression), and (c) applicable to school-age youth. The reviewers reconciled their screening decisions, which were in agreement (before reconciliation) for 93 percent of the programs. In addition, programs on eight lists classified as "promising" were screened out because the research on them was insufficient. The screening process reduced the number of programs from 317 to 89. (See Exhibit B-2 for a listing of the 89 programs, and B-3 for the final status of those programs after the review.)

Figure B-1. Summary of research review

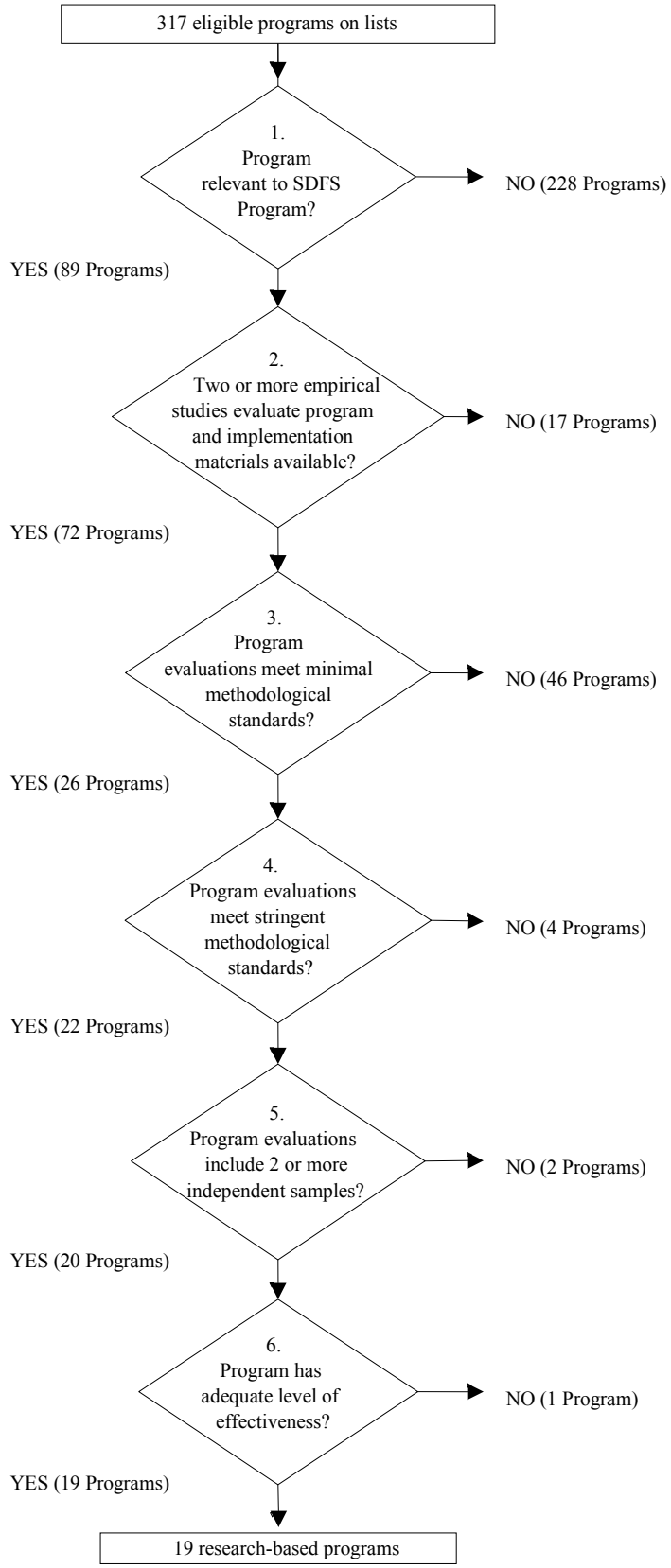

Exhibit B-1. Lists of promising and effective programs*

1. American Youth Policy Forum

Mendel, R.A. (2001). *Less Hype, More Help: Reducing Juvenile Crime, What Works – and What Doesn't.* Washington DC: American Youth Policy Forum. (http://www.aypf.org/)

2. Blueprints for Violence Prevention**

Elliott, D.S. (Series Editor) (1997). *Blueprints for Violence Prevention (Vols. 1–11).* Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado. (http://www.colorado.edu/cspv/blueprints/index.html)

3. Center for Mental Health Services**

Greenberg, M.T., Domitrovich, C., and Bumbarger, B. (1999). *Preventing Mental Disorders in School-Aged Children: A Review of the Effectiveness of Prevention Programs.* State College, PA: Prevention Research Center for the Promotion of Human Development, College of Health and Human Development, Pennsylvania State University. (http://www.prevention.psu.edu/)

4. Center for Substance Abuse Prevention (CSAP) **

Center for Substance Abuse Prevention (CSAP), Dept. of Health and Human Services, National Registry of Effective Programs (NREP). (modelprograms.samhsa.gov)

5. U.S. Department of Education Safe Schools**

(http://www.ed.gov/admins/lead/safety/exemplary01/panel.html)

6. Communities That Care—Developmental Research and Programs

Posey, R., Wong, S., Catalano, R., Hawkins, D., Dusenbury, L., and Chappell, P. (2000). *Communities That Care Prevention Strategies: A Research Guide to What Works.* Seattle, WA: Developmental Research and Programs, Inc. (http://www.channing-bete .com/positiveyouth/pages/CTC/prevention_strategies.html)

7. Mihalic and Aultman-Bettridge**

Mihalic, S., and Aultman-Bettridge, T. (2002). A Guide to Effective School-Based Prevention Programs. In W.L. Tulk (ed.), *Policing and School Crime*. Englewood Cliffs, NJ: Prentice Hall Publishers.

8. National Institute on Drug Abuse (NIDA)

National Institute on Drug Abuse**.** (1997). *Preventing Drug Use among Children and Adolescents: A Research-Based Guide* (NCADI # PHD734). Washington, DC: National Clearinghouse for Alcohol and Drug Information (NCADI). (http://165.112.78.61/Prevention/Prevopen.html)

9. Sherman et al.

Sherman et al. (1997). *Preventing Crime: What Works, What Doesn't, What's Promising* (NCJ 165366). Washington, DC: U.S. Department of Justice, Office of Justice Programs. (http://www.ncjrs.org/works/ or http://www.preventingcrime.org/)

10. Strengthening America's Families**

(http://www.strengtheningfamilies.org/)

11. Surgeon General's Report***

U.S. Department of Health and Human Services (2001). *Youth Violence: A Report of the Surgeon General.* Rockville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Injury Prevention and Control; Substance Abuse and Mental Health Services Administration, Center for Mental Health Services; and National Institutes of Health, National Institute of Mental Health. (http://www.surgeongeneral.gov/library/youthviolence/)

12. Title V (OJJDP) **

Title V. Training and Technical Assistance Programs for State and Local Governments: Effective and Promising Programs Guide. Washington DC: Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice.

* Mihalec, S. (2003). Matrix of programs as identified by various federal and private agencies. Retrieved from http://www.colorado.edu/cspv/blueprints/matrix/overview.html on Oct. 15, 2003.

** Promising programs on this list were dropped from consideration at step one.

*** Promising 2 programs and model 2 programs on this list were dropped from consideration at step one.

Exhibit B-2. Prevention programs aligned with the mission of the Safe and Drug-Free Schools and Communities Program

1. Across Ages
2. Adolescent Alcohol Prevention Trial (AAPT)/All Stars
3. Adolescent Transitions Program
4. Aggression Replacement Training
5. Alcohol Misuse Prevention
6. Al's Pals: Kids Making Healthy Choices
7. Anger Coping Program
8. Athletes Training and Learning to Avoid Steroids (ATLAS)
9. Behaviorally-Based Prevention Program
10. Bicultural Competence Skills Approach
11. Brainpower Program (Attributional Intervention)
12. Bullying Prevention Program (BPP) (also known as Intervention Campaign Against Bully/Victim Problems)
13. CAPSLE
14. CASASTART
15. Chicago Child-Parent Center and Expansion Program (CPC)
16. Child Development Project
17. Children of Divorce Intervention Program
18. Classroom-Centered (CC) and Family-School Partnership (FSP) intervention
19. Comer School Development Program
20. Consistency Management and Cooperative Discipline (CMDC)
21. Coping Power
22. Earlscourt Social Skills Group Program
23. Early Risers
24. Families and Schools Together
25. FAST Track
26. First Step to Success
27. Friendly PEERsuasion
28. Gang Prevention Curricula
29. Gang Resistance Education and Training (G.R.E.A.T.)
30. Good Behavior Game/Baltimore Mastery Learning
31. Growing Healthy
32. Guiding Good Choices (also known as Preparing for the Drug Free Years)
33. Healthy for Life
34. I Can Problem Solve (Interpersonal Cognitive Problem Solving)
35. Improving Social Awareness-Social Problem Solving
36. Incredible Years
37. Keep A Clear Mind (KACM)
38. Keeping It Real
39. Know Your Body
40. Leadership and Resiliency (LRP)
41. Life Skills Training
42. Linking the Interests of Families and Teachers (LIFT)
43. Lions-Quest Skills for Adolescence (also known as Skills for Adolescence)
44. Metropolitan Area Child Study
45. Michigan Model for Comprehensive School Health Education
46. Midwestern Prevention (Project STAR)
47. Montreal Longitudinal Experimental Study (Preventive Treatment Program)
48. Native American Prevention Project Against AIDS/Substance Abuse (NAPPASA)
49. PeaceBuilders
50. Peaceful Conflict Resolution and Violence Prevention Curriculum
51. Peer Coping Skills Training
52. Peers Making Peace
53. Positive Action
54. Positive Adolescent Choices Training (PACT)
55. Positive Youth Development Program
56. Proactive Classroom Management
57. Project ACHIEVE
58. Project ALERT
59. Project Northland
60. Project PATHE (Positive Action Through Holistic Education; also known as Promoting Action Through Holistic Education)
61. Project SUCCESS
62. Promoting Alternative Thinking Strategies (PATHS)
63. Protecting You/Protecting Me
64. Reconnecting Youth Program
65. Resolving Conflicts Creatively
66. Responding in Peaceful and Positive Ways (RIPP) / Richmond Youth Against Violence Project: Responding in Peaceful and Positive Ways (RIPP)
67. Rural Education Achievement Project
68. Safe Dates
69. School Safety Program
70. School Violence Prevention Demonstration Program
71. School-based Smoking Prevention Program
72. Schools and Families Education Children (SAFE Children)
73. Second Step
74. Sembrando Salud
75. Skills, Opportunities, and Recognition (SOAR) (also known as Seattle Social Development Program)
76. SMART Leaders
77. SMART Team
78. Social Competence Promotion Program for Young Adolescents (SCPPYA)
79. Social Relations Program
80. Socio-moral Reasoning Development Program
81. Strengthening Families Program
82. Student Training Through Urban Strategies (STATUS)
83. Students Helping Others Understand Tobacco (SHOUT)
84. Think First
85. Too Good For Drugs (TGFD)
86. Toward No Drug Abuse (TND)
87. Towards No Tobacco Use (TNT)
88. Washington (DC) Community Violence Prevention Program
89. Woodrock Youth Development Project

Exhibit B-3. Status of programs that remained after initial screening for relevance to study

**	1.	Across Ages
	2.	Adolescent Alcohol Prevention Trial (AAPT)/All Stars
	3.	Adolescent Transitions Program
	4.	Aggression Replacement Training
	5.	Alcohol Misuse Prevention
*	6.	Al's Pals: Kids Making Healthy Choices
	7.	Anger Coping Program
**	8.	Athletes Training and Learning to Avoid Steroids (ATLAS)
**	9.	Behaviorally-Based Prevention Program
*	10.	Bicultural Competence Skills Approach
	11.	Brainpower Program (Attributional Intervention)
***	12.	Bullying Prevention Program (BPP) (also known as Intervention Campaign Against Bully/Victim Problems)
**	13.	CAPSLE
**	14.	CASASTART
**	15.	Chicago Child-Parent Center and Expansion Program (CPC)
	16.	Child Development Project
**	17.	Children of Divorce Intervention Program
**	18.	Classroom-Centered (CC) and Family-School Partnership (FSP) intervention
**	19.	Comer School Development Program
**	20.	Consistency Management & Cooperative Discipline (CMDC)
**	21.	Coping Power
	22.	Earlscourt Social Skills Group Program
	23.	Early Risers
**	24.	Families and Schools Together
**	25.	FAST Track

___ Evaluated program
* Eliminated at Step 2
** Eliminated at Step 3
*** Eliminated at Step 4
△ Eliminated at Step 5
§ Eliminated at Step 6

Exhibit B-3. Status of programs that remained after initial screening for relevance to study—continued

**	26.	First Step to Success
**	27.	Friendly PEERsuasion
**	28.	Gang Prevention Curricula
§	29.	Gang Resistance Education and Training (G.R.E.A.T.)
**	30.	Good Behavior Game/Baltimore Mastery Learning
**	31.	Growing Healthy
△	32.	Guiding Good Choices (also known as Preparing for the Drug Free Years)
**	33.	Healthy for Life
***	34.	I Can Problem Solve (Interpersonal Cognitive Problem Solving)
**	35.	Improving Social Awareness-Social Problem Solving
	36.	Incredible Years
*	37.	Keep A Clear Mind (KACM)
**	38.	Keeping It Real
	39.	Know Your Body
*	40.	Leadership and Resiliency (LRP)
	41.	Life Skills Training
**	42.	Linking the Interests of Families and Teachers (LIFT)
**	43.	Lions-Quest Skills for Adolescence (also known as Skills for Adolescence)
**	44.	Metropolitan Area Child Study
**	45.	Michigan Model for Comprehensive School Health Education
	46.	Midwestern Prevention (Project STAR)
**	47.	Montreal Longitudinal Experimental Study (Preventive Treatment Program)
*	48.	Native American Prevention Project Against AIDS/Substance Abuse (NAPPASA)
**	49.	PeaceBuilders
*	50.	Peaceful Conflict Resolution and Violence Prevention Curriculum
**	51.	Peer Coping Skills Training
*	52.	Peers Making Peace

___ Evaluated program
* Eliminated at Step 2
** Eliminated at Step 3
*** Eliminated at Step 4
△ Eliminated at Step 5
§ Eliminated at Step 6

Exhibit B-3. Status of programs that remained after initial screening for relevance to study—continued

53. Positive Action
* 54. Positive Adolescent Choices Training (PACT)
* 55. Positive Youth Development Program
* 56. Proactive Classroom Management
** 57. Project ACHIEVE
58. Project ALERT
** 59. Project Northland
** 60. Project PATHE (Positive Action Through Holistic Education; also known as Promoting Action Through Holistic Education)
** 61. Project SUCCESS
62. Promoting Alternative Thinking Strategies (PATHS)
* 63. Protecting You/Protecting Me
** 64. Reconnecting Youth Program
** 65. Resolving Conflicts Creatively
66. Responding in Peaceful and Positive Ways (RIPP) / Richmond Youth Against Violence Project: Responding in Peaceful and Positive Ways (RIPP)
* 67. Rural Education Achievement Project
** 68. Safe Dates
** 69. School Safety Program
* 70. School Violence Prevention Demonstration Program
** 71. School-based Smoking Prevention Program
** 72. Schools and Families Education Children (SAFE Children)
73. Second Step
** 74. Sembrando Salud
* 75. Skills, Opportunities, and Recognition (SOAR) (also known as Seattle Social Development Program)
** 76. SMART Leaders
** 77. SMART Team

___ Evaluated program
* Eliminated at Step 2
** Eliminated at Step 3
*** Eliminated at Step 4
△ Eliminated at Step 5
§ Eliminated at Step 6

Exhibit B-3. Status of programs that remained after initial screening for relevance to study—continued

** 78. Social Competence Promotion Program for Young Adolescents (SCPPYA)

* 79. Social Relations Program

* 80. Socio-moral Reasoning Development Program

△ 81. Strengthening Families Program

** 82. Student Training Through Urban Strategies (STATUS)

** 83. Students Helping Others Understand Tobacco (SHOUT)

** 84. Think First

** 85. Too Good For Drugs (TGFD)

86. Toward No Drug Abuse (TND)

** 87. Towards No Tobacco Use (TNT)

* 88. Washington (DC) Community Violence Prevention Program

*** 89. Woodrock Youth Development Project

___ Evaluated program
* Eliminated at Step 2
** Eliminated at Step 3
*** Eliminated at Step 4
△ Eliminated at Step 5
§ Eliminated at Step 6

Assessing the quality of evidence on programs and patterns of program-related effects. For steps two through six of the review process, research literature on the programs that survived the screening process (step one) was gathered and reviewed. This review focused on assessing the quality of research evidence (steps two through five); and, for research that met study standards of quality, assessing the patterns of program-related effects for the programs involved (step 6). It entailed the following steps.

Step two—review abstracts and search for materials to support dissemination. After completing extensive automated searches for the published literature on the 89 programs, abstracts for over 2,100 reports on studies were reviewed for general relevance to the study; if fewer than two abstracts were found on empirical research on a program, the program was eliminated. In addition, whether programs had mechanisms in place to support the widespread dissemination of the programs was determined (e.g., training materials were available); if no such mechanisms for given program were found, it was eliminated. This step reduced the number of programs from 89 to 72.

Step three—review research on programs against minimal standards. The programs were further screened on whether they had studies on them that met at least minimally acceptable methodological standards. Using a standardized form, reviewers completed the following set of tasks for each program.

- First, the reviewers examined whether each study reported quantitative results on behavioral outcomes for the program of interest; these behavioral outcomes had to pertain to youth substance abuse or aggressive or criminal behavior (i.e., attitudes and minor disruptive behaviors were excluded).

- Second, the reviewers recorded whether or not any statistically significant results ($p < 0.05$) favored the program on any of the behavioral outcomes.

- Third, the reviewers ascertained whether an "acceptable" research design was used: randomized control trial, pretest and posttest with comparison group, regression discontinuity, or interrupted time series.

- Fourth, the reviewers looked across the studies that had passed the step three criteria and decided whether at least two independent samples were studied.

A program passed step three if at least two studies on it yielded significant results on behavioral outcomes, used one of the four acceptable designs, and used at least two unique samples. This step reduced the number of programs from 72 to 26.

Step four—review research on programs against stringent standards. Following completion of a training session, two staff with strong methodology skills independently reviewed each study on the surviving programs that used behavioral outcomes and an acceptable design. The reviewers used a standardized form to evaluate studies on several methodological criteria, some but not all of which are based on the Standards of Evidence developed by the Society for Prevention Research (Flay et al., 2005). For reports that described multiple substudies, reviewers examined the one substudy that best represented the program of interest, used the strongest design, and included the first post-program follow-up time point. To pass step four, a study had to meet each of the following criteria.

- Construct validity
  - Intervention as planned was similar to the program model,
  - Intervention was implemented as planned,
  - No expectancy, novelty, disruption effect occurred for the treatment group, and
  - Outcome was aligned with the intervention.
- Internal validity
  - Minimal differential attrition occurred,
  - Quasi-experimental—adequate group equating procedures were used,
  - No local history effect occurred, and
  - Other contaminants were unlikely.
- External validity or generalizability of findings
  - Sample represented the population of interest (English-speaking North American school-age youths at the level of risk targeted by the program).

The reviewers reconciled their ratings on each of these and other criteria, which were in agreement (before reconciliation) for approximately 90 percent of the ratings. The results of step four reduced the number of programs from 26 to 22.

Step five—check for independent samples. For each of the programs that passed step four, whether the surviving studies on it used at least two independent samples of participants was examined (to help ensure that the study results have been replicated). This step reduced the number of programs from 22 to 20.

Step six—examine patterns of program-related effects. For programs supported by evidence that met study standards of quality (i.e., passed steps one through five), an in-depth review of the level of effectiveness indicated by the studies was conducted. For each program, the pattern of findings on behavioral outcomes related to youth substance abuse and aggressive and criminal behavior was examined. This review, which was both quantitative and qualitative, considered the statistically significant ($p < 0.05$) program-related findings reported, as well as the factors that may have influenced those findings (e.g., sample size) and how those findings varied by study and type of outcome. A program was recommended for the list of Research-Based Programs if the studies on it demonstrated a positive pattern of program-related findings on the outcomes of interest. This step reduced the number of programs from 20 to 19.

On Jan. 13, 2006, an expert panel was convened to critique the research review methodology and to examine the appropriateness of the programs selected for inclusion and suggest other programs that may have been missed through this process. The panel members have expertise in the following four areas: (a) reviewing quality of evidence, (b) meta-analysis, (c) measuring implementation, and (d) school-based prevention. (See Exhibit B-4 for a list of panel members and their qualifications.)

Exhibit B-4. Expert panel members

Brian Flay, Ph.D., University of Illinois at Chicago. Dr. Flay led efforts to develop standards on the quality of evidence for the Society on Prevention Research; these standards are closely aligned with the What Works Clearinghouse approach. He is also an expert on prevention programming.

Harris Cooper, Ph.D., Duke University. Dr. Cooper is a co-developer of What Works Clearinghouse procedures and instruments for reviewing research publications. He is also an expert on meta-analytic methods.

Fred Springer, Ph.D., EMT Associates. Dr. Springer led the Center for Substance Abuse Prevention (CSAP) National Cross-site Evaluation Study of High Risk Youth Programs, which used innovative approaches to measuring implementation.

Mike Furlong, Ph.D., University of California, Santa Barbara. Dr. Furlong has conducted research and developed programs related to school safety, school violence, and anger management in youth.

Judy Thorne, Ph.D., has conducted research in the areas of substance abuse prevention and school-based drug prevention programming. She is an expert on school-based prevention.

The list of Research-based Programs for the Prevalence Study, which reflects input from the expert panel, includes the 19 programs that met criteria for the study. In addition, to help reconcile the study list of research-based programs with another list that became available after the research review was underway (Helping America's Youth), the U.S. Department of Education requested that two programs be included with the 19 research-based programs in the Prevalence Study analyses on research-based programs in schools.[31] (See Exhibit B-5 for the final list of Research-based Programs.) The analyses for the Fidelity Study include only the 19 programs that met criteria for the study. The programs on these lists are likely to differ from those on other lists for a variety of reasons, including the study-specific review criteria that were used.

Practices. To develop a list of research-based practices, a review of the evidence provided by meta-analyses was conducted, and the list of research-based practices developed by Gottfredson et al. (2000) was used. With regard to the latter list, the two principal investigators for the Gottfredson et al. study independently identified practices for a given type of program (e.g., curriculum, instruction, or training program) based on their knowledge of the research literature. They were in high agreement on their initial judgments on the practices. To reconcile discrepancies in judgments, they discussed the discrepancies and referred to the supporting evidence. Their reconciled list of practices is reflected in the Fidelity Study.

[31] The two programs are Skills, Opportunities, and Recognition (SOAR) and Too Good for Drugs. SOAR had been excluded from the study list of research-based programs because it lacked implementation materials. Too Good for Drugs might have been included on the list had the review considered a publication that was deemed out of scope (because it was a technical report with limited circulation rather than a published document).

Exhibit B-5. Descriptions of effective programs resulting from research synthesis

1. Adolescent Alcohol Prevention Trial (AAPT)/All Stars

All Stars is a school- or community-based program designed prevent or delay the onset of high-risk behaviors (e.g., drug use, violence, and premature sexual activity) in middle schools youth (ages 11-14). This intervention is designed to help adolescents develop qualities that will motivate them to avoid drug use and high-risk behaviors, reduce the use of gateway drugs, develop meaningful relationships, and develop positive characters and lifestyles.

2. Aggression Replacement Training

Aggression Replacement Training is designed to teach adolescents to understand and replace aggression and antisocial behavior with positive alternatives. The program's three-part approach includes training in prosocial skills, anger control, and moral reasoning.

3. Alcohol Misuse Prevention (AMP)

Alcohol Misuse Prevention is a school-based alcohol prevention program designed to prevent the misuse of alcohol by adolescents. This curriculum presents basic information on alcohol and its effects, and focuses on making safe and informed decisions. Students practice building resistance skills through role plays.

4. Anger Coping Program

The Anger Coping Program is a school-based intervention for youths (ages 8–14) who have been teacher-identified as aggressive and disruptive. Groups of five to seven students meet once-per-week for 45 to 60 minutes for 18 sessions. The program is based on a social-cognitive model of anger and is designed to reduce future conduct problems, delinquency, and substance abuse.

5. Adolescent Transitions Program

The Adolescent Transitions Program is a parent training program developed as a selected intervention for at-risk early adolescents. The parent-focused curriculum is based on family management skills of encouragement, limit setting and supervision, problem solving, and improved family relationship and communication patterns. These skills follow a step-wise approach toward effective parenting skills and strategies for maintaining change. The long-term goals of the program are to arrest the development of teen antisocial behaviors and drug experimentation; the intermediate goals of the program are to improve parent family management and communication skills.

Exhibit B-5. Descriptions of effective programs resulting from research synthesis—continued

6. Brainpower (Attributional Intervention)

The BrainPower Program is a 12-lesson school-based intervention. Originally implemented with African-American elementary school students, this program is designed to change hostile attributional biases such as verbal or nonverbal behaviors that are misread as hostile or threatening by aggressive children. The curriculum utilizes a variety of strategies including role-play, discussion of personal experiences, and brainstorming.

7. Child Development Project

The Child Development Project is a school-based intervention aimed at reducing early use of alcohol and marijuana and improving violence-related behavior. The program was developed for elementary school youths (ages 5–12) and includes a buddy system of older and younger students, activities for students to complete at home with parents or caregivers, activities involving relatives at school, class meetings, and literacy-building activities.

8. Earlscourt Social Skills Group Program

The Earlscourt Social Skills Group Program is school-based intervention designed to improve the self-control and social skills of aggressive, noncompliant children (ages 6–12). Eight basic skills are taught: problem solving, knowing your feelings, listening, following instructions, joining in, using self-control, responding to teasing, and keeping out of fights.

9. Early Risers

Early Risers is a multi-component, high-intensity, competency-enhancement program that targets elementary school children (ages 6–12) at high risk for early development of conduct problems, including substance use. The program has a child-focused component and a parent-focused component. A family advocate visits the child's school, consults with teachers, and mentors the student.

10. Incredible Years

The Incredible Years comprises three comprehensive, multifaceted, and developmentally based curricula for parents, teachers, and children (ages 2–8). The program is designed to promote emotional and social competence; and to prevent, reduce, and treat aggressive, defiant, oppositional, and impulsive behaviors in young children.

Exhibit B-5. Descriptions of effective programs resulting from research synthesis—continued

11. Know Your Body

Know Your Body is a comprehensive school health promotion program for kindergarten through sixth grade. Children are taught the connections between smoking-related decisions and self-image, values, anxiety, and stress, as well as skills in stress management, decision-making, communication, and assertiveness. The program has five components: skills-based health education curriculum and teacher or coordinator training are core components; biomedical screening, extracurricular activities, and program evaluation are enhancements.

12. Life Skills Training

LifeSkills Training is a school-based substance abuse and violence prevention program for upper elementary and middle school students (ages 11–14). Students are taught personal self-management skills, general social skills, drug resistance skills, adaptive coping strategies, assertiveness, and decision-making by either adults or peer leaders.

13. Midwestern Prevention (Project Star)

Project STAR is a comprehensive, community-based drug abuse intervention program that uses school, mass media, parent education, community organization, and health policy programming to prevent and reduce tobacco, alcohol, marijuana, and other drug use by adolescents. The program offers a series of classroom-based sessions for the school program during middle school that continues with the parent, media, community, and policy components.

14. Positive Action

Positive Action is a comprehensive program for children and adolescents ages 5 to 18 years. This program is intended to help students to learn and practice positive thoughts, actions, and feelings; decrease drug, alcohol, and tobacco use, disruptive behaviors, truancy, suspensions, and dropouts; and increase academic achievement, self-esteem, social development, positive behaviors, self-responsibility, and character development. It includes school, family, and community components that work together or stand alone.

15. Project ALERT

Project ALERT is a school-based drug prevention program for middle school students (ages 11–14). The curriculum focuses on the substances that adolescents are most likely to use (e.g., alcohol, tobacco, marijuana, and inhalants). Project ALERT is intended to motivate adolescents not to use drugs by teaching them the skills and strategies needed to resist social pressures to use drugs, and to establish nondrug-using norms.

Exhibit B-5. Descriptions of effective programs resulting from research synthesis—continued

16. Promoting Alternative Thinking Strategies (PATHS)

PATHS is a curriculum-based program for elementary school aged children (ages 5–12) designed to facilitate the development of self-control, emotional awareness, and interpersonal problem-solving skills. The program is intended to reduce aggression and behavior problems while simultaneously enhancing emotional development and the educational process in the classroom.

17. Responding in Peaceful and Positive Ways (RIPP)/Richmond Youth against Violence Project: Responding in Peaceful and Positive Ways

RIPP is a school-based violence prevention program designed to provide middle schools students with conflict resolution strategies and skills. The program combines classroom instruction in problem solving with opportunities for peer mediation. RIPP promotes nonviolence by teaching students more effective ways of dealing with interpersonal conflicts, and by lowering the number of violent incidents in school settings.

18. Second Step

Second Step is a school-based social skills program for pre-school through junior high students (ages 4–14). It is designed to reduce impulsive, high-risk, and aggressive behaviors, and to increase children's social-emotional competence and other protective factors.

19. Toward No Drug Abuse (TND)

Project Toward No Drug Abuse (TND) is a school-based interactive prevention program designed to help high school youth (ages 14–19) resist substance use. The program consists of twelve 40- to 50-minute lessons. The curriculum includes motivational activities, social skills training, and decision-making components that are delivered through group discussions, games, role-playing exercise, videos, and student worksheets.

As background on the review of meta-analyses, in meta-analysis, an estimate of effectiveness is typically indexed by Hedges' *d*, a standardized estimate that reflects the difference in central tendency between the group receiving treatment and the group not receiving treatment after accounting for the variation in outcome across the groups. This standardized estimate, or "effect size" is expressed as a *z-score*, which means that it is recorded on a scale where 0 equals no difference between the groups, 1 or -1 equals about a 40 percent difference between the groups, and 2 equals about a 70 percent difference between the groups (Rosenthal, 1994). In the world of treatment effectiveness research, experience has suggested that effect sizes in the range of $d = 0.20 - 0.49$ are generally considered small, effect sizes $0.50 - 0.79$ are considered medium, while any effect size 0.80 or larger is considered a big effect (Cohen, 1988).[32]

The key words "meta-analysis AND school" were used to identify a total of 44 meta-analyses from several online literature search vehicles. Seventeen of these meta-analyses were retrieved and reviewed for relevant content. All estimates of effectiveness that were based on posttest comparisons of an intervention and control or comparison group were extracted to a machine readable database indicating the magnitude of the effect size, the number of estimates contributing to the effect size, and the 95 percent confidence interval for the estimate. In addition to these numeric values, the principle being summarized and the outcome being tested by the effect size were recorded. This allowed sorting and organizing of the evidence across disparate meta-analysis to make judgments on which principles were associated with stronger effects, and which with weaker effects, across various outcomes.

A total of 220 estimates of the effectiveness of various components of intervention were coded from 7 of the 14 meta-analyses. The principles tested ranged from type of leadership, to theoretical orientation of the intervention, to length of followup to publication source of the original research. Outcomes included substance use, delinquency, problem or antisocial behavior, and dropout or truancy.

The review applied three criteria to establish which practices were associated with the greatest impact on intervention recipients. The first was effect size: Estimates had to be greater than $d = 0.20$ (Cohen's [1988] baseline for a small effect size). According to Rosenthal, an effect size of $d = 0.20$ is the equivalent to about a 10 percent standardized difference between

[32] Note however, that these values are arbitrary and were proposed merely as a guide for thinking about the relative magnitude and meaning of the effect size observed.

intervention and control or comparison groups. The second criterion pertained to the "density" of evidence for a mean estimate of effectiveness. While no generally accepted standard at which signifies "enough" data exists, for this study, a standard of at least six primary studies contributing estimates was used. The third criterion was consistency of estimates across outcome domains and meta-analyses. All other things being equal, greater consistency suggests that the results are relatively more robust and replicable.

Thirty-two estimates met or exceeded the effect size criterion. The database was then sorted to assess whether results for the constructs represented by these 32 estimates were positive and generally met or exceeded the study criterion of $d = 0.20$. In this second step, several principles were dropped as results were not consistent across outcomes or meta-analyses. Seven principles of effectiveness survived this process and were identified as effective principles based on the procedures followed in this review. The principles identified as effective are as follows:

- Having a clinician leader (Tobler et al., 2000);
- Having a peer leader or facilitator (Bangert-Drowns, 1988; Gottfredson and Wilson, 2003; Tobler et al., 2000);
- System-wide change (Tobler et al., 2000);
- Cognitive behavioral, behavioral modeling, or behavioral modification (Wilson, Gottfredson, and Najaka, 2001);
- Increase self-esteem and self reliance, decrease alienation, improve decision and interpersonal skills (Bruvold, 1993);
- Social norms, reduce alienation, increase self-esteem (Bruvold, 1993); and
- Recognize and resist social pressure, immediate social and physical consequences of use (Bruvold, 1993).

APPENDIX C
EFFORTS TO COLLECT DATA ON PROGRAM FUNDING FROM SCHOOL PERSONNEL

The Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime sought to answer research questions on the prevalence and implementation fidelity of research-based prevention programs in public elementary and secondary schools, including those programs that received funding[33] from the Safe and Drug-Free Schools and Communities Act (SDFSCA) Program. Answering these questions required discerning whether specific school-level programs received any SDFSCA funding. Despite the best efforts of the research team, the responses gathered from schools on funding raised questions about their validity.

In this appendix, we discuss the issues surrounding survey data collection on funding for school-level activities. We summarize the approach taken by the study and the concerns about the results obtained. In addition, we present the results on SDFSCA funding pertaining to the adoption and implementation of research-based prevention programs in schools. Because of the concerns discussed, we urge caution in interpreting these results and drawing conclusions based on them. We also summarize lessons learned on gathering data on funding that can be applied to subsequent survey efforts.

Introduction and Background

Education researchers often struggle with how best to gather reliable information on aspects of funding for school-level activities. Part of the challenge is that funding decisions are typically made at the district level, while the activities occur at the school level. Hence, collecting this type of information typically requires the exchange of information between the district and school levels, which can be very difficult to achieve, especially for a large survey.

The Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime attempted to overcome this challenge by surveying school administrators (in the Prevention Program Survey) on whether specific programs in their schools received SDFSCA funding, and encouraging them to coordinate with district officials on this

[33] SDFSCA Program funding was eliminated beginning in FY 2010.

information. For each prevention program that the administrator identified as operating during the 2004–05 school year, he or she was asked to indicate whether the school received any SDFSCA funding. As mentioned, the instructions in the self-administered questionnaire encouraged the respondents to consult with their district prevention coordinators on funding and respond accordingly; also, we notified the district prevention coordinators that school staff may be contacting them on this issue. The information gathered on funding was used in analyses on whether SDFSCA Program funding was used to support research-based programs and on the association between SDFSCA Program funding and the implementation fidelity of research-based curriculum programs.

Unfortunately, the responses from school-level respondents in the Prevention Program Survey on funding were disappointing. Although the overall response rate for this survey was a respectable 86 percent, item nonresponse (missing or don't know) on SDFSCA Program funding was approximately 30.0 percent across the prevention programs reported; it was substantially higher for some specific programs. This indicates that many respondents were unclear on whether their programs received SDFSCA Program funding. Hence, the research team and the U.S. Department of Education agreed that the results on funding should be interpreted with caution, and moved from the main body of the report to this appendix.

Findings on the Prevalence of Programs Funded by the SDFSCA Program, Including Research-based Programs

These findings are presented on programs overall and research-based programs.

Programs overall funded by SDFSCA program. Schools reported that over 98,000 programs were supported by the U.S. Department of Education's SDFSCA Program during the 2004–05 school year (see Table C-1). Although schools implemented a large number of programs supported by the SDFSCA Program in 2004–05, these represented only 17.9 percent of programs overall.

In addition, programs are similar in regard to funding across school instructional levels. High school prevention programs appeared somewhat more likely to be supported by the SDFSCA Program funding than middle or elementary school programs, but the differences are slight. A test of statistical significance could not be conducted for these results.

Table C-1. Prevention programs by whether they received funding from SDFSCA Program and instructional level: 2004–05

Received SDFSCA funding	Instructional level							
	Elementary		Middle		High		All schools	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Yes	17.3	55,258	18.6	20,035	19.8	20,806	17.9	98,525
No	82.7	263,393	81.4	87,478	80.2	84,093	82.1	450,454
Total	100.0	318,651	100.0	107,513	100.0	104,899	100.0	548,979

NOTE: Tests of significance could not be performed. Estimates are based on responses from 3,451 schools. Schools that could either not specify the source of program funds or did not respond to the question were omitted from the analysis.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006; U.S. Department of Education, National Center for Education Statistics, "Common Core of Data," 2003–04.

Research-based programs funded by SDFSCA program. Most prevention programs supported by the SDFSCA Program were not research based. Approximately 12 percent of programs that received funds from the SDFSCA Program in 2004–05 had firm research grounding (see Table C-2). However, examining the data from the perspective of programs rather than funding source reveals that research-based programs more frequently received funding from the SDFSCA Program than non-research-based programs. Specifically, 32.4 percent of research-based programs received funds from the SDFSCA Program compared to 17.0 percent of other programs.

Table C-2. Prevention programs by whether research-based and received funding from SDFSCA program: 2004–05

Research-based program	Received funding from SDFSCA program					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Yes	11.9	11,750	5.4	24,503	6.6	36,253
No	88.1	87,360	94.6	427,773	93.4	515,133
Total	100.0	99,110	100.0	452,276	100.0	551,386

NOTE: A test of statistical significance could not be performed for this table. Estimates are based on responses from 3,468 schools. Schools that could not specify the source of program funds or did not respond to the question were omitted from the analysis.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006.

Conclusions. In considering the findings on SDFSCA Program funding for programs, one should keep in mind the context for them. First, although the legislation authorized the use of SDFSCA Program funds for a broad range of activities, the study focused on programs intended to prevent youth ATOD use and school crime, which is a subset of those

activities. Also, under the statute, local school districts could have applied to state education agencies for a waiver of the requirement of using SDFSCA Program funds only for research-based programs; those non-research-based programs must have been innovative programs that demonstrated a substantial likelihood of success. Finally, because the list of research-based programs prepared for the study was intended to reflect the results of current rigorous research on programs, it excludes some programs that appear on other lists of promising and effective programs either because there was a lack of research evidence or that the program's scope was dissimilar.[34] As a result, schools might have intended to apply the "Principles of Effectiveness" standard of implementing a research-based program but did not achieve it based on the study's criteria.

With regard to the question on programs receiving funding from the SDFSCA Program, the Prevalence Study found that only 11.9 percent of the programs that received funding from the SDFSCA Program were research based. Although this proportion is higher than for programs funded by other sources (5.4 percent), it is quite low despite the U.S. Department of Education's expectation that programs with SDFSCA Program funding be research based. While one can debate the extent to which the low proportion is due to differences between the Prevalence Study and other sources on the definition of research-based programs, the fact remains that nearly 90 percent of SDFSCA-funded programs lacked adequate empirical support for their effectiveness in preventing or reducing youth substance abuse and school crime.

Findings on the Implementation Fidelity of Research-based Curriculum Programs, Including Those Funded by the SDFSCA Program

These findings are on the relationship between implementation fidelity and receipt of Safe and Drug-Free Schools and Communities (SDFSCA) Program funding. They describe the receipt of SDFSCA Programs funding for the programs included in the fidelity analyses, and how those programs fared on program-specific and generic fidelity standards.

Research-based programs included in the fidelity analyses. Among the programs for which funding information was available, 40.1 percent of research-based curriculum programs implemented during the 2004–05 school year received funding from the Department of

[34] In addition, the list excludes programs that have not been adequately evaluated, including programs that are relatively more difficult to evaluate. Rather than indicating that these programs are ineffective, insufficient evidence exists on their effectiveness.

Education's SDFSCA Program (see Table C-3). Second Step programs account for 42.0 percent of those that received SDFSCA support, followed by Life Skills Training programs (23.6 percent) and Project Alert programs (16.6 percent). All other curriculum programs account for 17.8 percent of the research-based curriculum programs that received SDFSCA Program funding.

Table C-3. Research-based curriculum programs in fidelity analyses by whether received funding from SDFSCA Program: 2004–05

Program	Received funding from SDFSCA Program[a]					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Life Skills Training.....	23.6	1,372	24.2	2,100	23.9	3,472
Project Alert.....	16.6	963	9.6	836	12.4	1,799
Second Step	42.0	2,447	40.4	3,511	41.1	5,958
All other programs.....	17.8	1,037	25.8	2,238	22.6	3,275
Total.....	100.0	5,819	100.0	8,686	100.0	14,505

[a] Estimates are based on responses on 565 programs. $x^2 = 7.85$, $p < .05$.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006; "Provider Survey," 2007.

Achievement of program-specific standards. For the programs on which funding information was available, the proportion passing on program-specific standards was equivalent for those curriculum programs that received funding from the SDFSCA Program and those that did not receive funding from that source. For example, 90.5 percent of the programs that received funding passed on topics covered versus 92.7 percent of programs without that funding (see Table C-4). Examined from another perspective, 39.9 percent of the programs passing on topics covered received funding from the SDFSCA Program, as did 46.8 percent of the programs failing on that standard. The relationship between passing on a program-specific standard and receiving funding from the SDFSCA Program was not statistically significant.

Table C-4. Research-based curriculum programs passing on program-specific standards by whether they received funding from SDFSCA Program: 2004–05

Passed	Received funding from SDFSCA Program					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Topics covered						
Yes	90.5	5,213	92.7	7,852	91.8	13,065
No.....	9.5	545	7.3	620	8.2	1,165
Total	100.0	5,758	100.0	8,472	100.0	14,230
Number of lessons						
Yes	68.4	2,913	64.1	4,027	65.9	6,940
No.....	31.6	1,344	35.9	2,250	34.1	3,595
Total	100.0	4,258	100.0	6,277	100.0	10,535
Frequency of student participation						
Yes	62.8	3,405	66.8	5,001	65.1	8,407
No.....	37.2	2,018	33.2	2,488	34.9	4,506
Total	100.0	5,423	100.0	7,489	100.0	12,912
Targeting on risk level						
Yes	91.9	4,932	86.9	7,061	88.9	11,993
No.....	8.1	434	13.1	1,063	11.1	1,498
Total	100.0	5,366	100.0	8,124	100.0	13,491

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006; "Provider Survey," 2007.

Achievement of generic standards. Among programs for which funding information was available, the proportion passing on each generic fidelity standard was similar for those curriculum programs that received SDFSCA funding and those that did not receive funding from that source (see Table C-5). For example, 76.7 percent of the programs that received SDFSCA Program funding passed on instructional strategies versus 78.7 percent of programs without that funding. Similarly, 32.0 percent of the programs that received funding passed on rewards, recognition, and mastery assessment versus 34.3 percent of programs without that funding. The relationship between passing on a generic fidelity standard and receiving funding from the SDFSCA Program was not statistically significant.

Table C-5. Research-based curriculum programs passing on generic standards by whether they received funding from SDFSCA Program: 2004–05

Passed	Received funding from SDFSCA Program					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Instructional strategies						
Yes	76.7	4,414	78.7	6,644	77.9	11,058
No	23.3	1,344	21.3	1,794	22.1	3,138
Total	100.0	5,758	100.0	8,438	100.0	14,195
Rewards, recognition, and mastery assessment						
Yes	32.0	1,799	34.3	2,935	33.4	4,734
No	68.0	3,817	65.7	5,625	66.6	9,441
Total	100.0	5,615	100.0	8,560	100.0	14,175

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006, "Provider Survey," 2007.

Overall implementation fidelity. The proportion of programs implemented with adequate quality among programs for which funding information was available was similar for those curriculum programs that received SDFSCA Program funding and those that did not (see Table C-6). For example, 46.3 percent of the programs that received SDFSCA Program funding were implemented with adequate overall quality versus 47.9 percent of programs without that funding. Similarly, 48.9 percent of the programs that received some SDFSCA Program funding passed on 3 to 4 standards compared to 43.7 percent of programs that did not receive funding from the Program. The relationship between overall implementation fidelity and receiving funding from the SDFSCA Program was not statistically significant.

Table C-6. Overall implementation fidelity of research-based curriculum programs by whether they received funding from SDFSCA Program: 2004–05

Overall implementation fidelity	Received funding from SDFSCA Program					
	Yes		No		Total	
	Percent	Weighted number	Percent	Weighted number	Percent	Weighted number
Inadequate[a]	53.7	2,061	52.1	2,786	52.7	4,847
Adequate[b]	46.3	1,780	47.9	2,562	47.3	4,342
Total	100.0	3,841	100.0	5,348	100.0	9,188

[a] Overall implementation fidelity is defined as "inadequate" if the program did not pass on all program-specific standards.

[b] Overall implementation fidelity is defined as "adequate" if the program passed on all program-specific standards.

SOURCE: U.S. Department of Education, Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime, "Prevention Program Survey," 2006, "Provider Survey," 2007.

Conclusions. One might expect that programs that received funding from the U.S. Department of Education's SDFSCA Program were better implemented overall. However, this is

not the case. For programs on which funding information was available, the Fidelity Study found that receiving funding from the SDFSCA Program was unrelated to the implementation fidelity. That is, programs with Program funding performed as well as those without such funding. Hence, the findings on programs nationally largely apply to the programs that received funding from the SDFSCA Program. These findings have a caveat: Approximately one-third of providers did not report on funding. To the extent that those providers delivered curriculum programs that did receive funding from the SDFSCA Program and those programs outperformed the ones on which reports are available, the findings may understate the quality of the programs funded by the Program.

Lessons Learned

The results from the Study of the Implementation of Research-Based Programs to Prevent Youth Substance Abuse and School Crime has reinforced the difficulty of gathering reliable information from school-level survey respondents on funding sources. While the approach used by the study appears to have obtained valid data on funding for the majority of prevention programs, it left questions about too many of the programs. All in all, that approach, which relied on school-level respondents consulting with district staff, was inadequate for achieving the desired level of response on funding.

Going forward, when possible, alternative approaches should be used that can better facilitate the exchange of information between the district and school levels on the funding of school-level activities. One such approach entails building in the opportunity for district prevention coordinators to confirm whether funds are used for each of the prevention programs reported by school administrators and encouraging the district prevention coordinators to consult other district staff and records on the funding sources for specific programs. Doing so requires conducting the district survey after the school survey on prevention programs. If surveys are primarily Web-based, the exchange of the information between surveys is feasible and efficient.

The Department of Education's mission is to promote student achievement and preparation for global competitiveness by fostering educational excellence and ensuring equal access.

http://www.ed.gov