As a library, NLM provides access to scientific literature. Inclusion in an NLM database does not imply endorsement of, or agreement with, the contents by NLM or the National Institutes of Health. Learn more about our disclaimer.

HHS Public Access
Author manuscript
Peer-reviewed and accepted for publication
About author manuscripts | Submit a manuscript

Am J Prev Med. Author manuscript; available in PMC 2008 Aug 1.
Published in final edited form as:
Am J Prev Med. 2007 Aug; 33(2 Suppl): S130–S143.
doi: 10.1016/j.amepre.2007.04.011

PMCID: PMC2246021
NIHMSID: NIHMS28281
PMID: 17675014

School-Based Interventions for Aggressive and Disruptive Behavior: Update of a Meta-Analysis

Sandra Jo Wilson, PhD[1] and Mark W. Lipsey, PhD[1]

Abstract

Background

Research about the effectiveness of school-based psychosocial prevention programs for reducing aggressive and disruptive behavior was synthesized using meta-analysis. This work updated previous work by the authors and further investigated which program and student characteristics were associated with the most positive outcomes.

Methods

Two hundred forty-nine experimental and quasi-experimental studies of school-based programs with outcomes representing aggressive and/or disruptive behavior were obtained. Effect sizes and study characteristics were coded from these studies and analyzed.

Results

Positive overall intervention effects were found on aggressive and disruptive behavior and other relevant outcomes. The most common and most effective approaches were universal programs and targeted programs for selected/indicated children. The mean effect sizes for these types of programs represent a decrease in aggressive/disruptive behavior that is likely to be of practical significance to schools. Multi-component comprehensive programs did not show significant effects

and those for special schools or classrooms were marginal. Different treatment modalities (e.g., behavioral, cognitive, social skills) produced largely similar effects. Effects were larger for better implemented programs and those involving students at higher risk for aggressive behavior.

Conclusions

Schools seeking prevention programs may choose from a range of effective programs with some confidence that whatever they pick will be effective. Without the researcher involvement that characterizes the great majority of programs in this meta-analysis, schools might be well-advised to give priority to those that will be easiest to implement well in their settings.

Introduction

Schools are an important location for interventions to prevent or reduce aggressive behavior. They are a setting in which much interpersonal aggression among children occurs and the only setting with almost universal access to children. There are many prevention strategies from which school administrators might choose, including surveillance (e.g., metal detectors, security guards); deterrence (e.g., disciplinary rules, zero tolerance policies); and psychosocial programs. Over 75% of schools in one national sample reported using one or more of these prevention strategies to deal with behavior problems.[1] Other reports similarly indicate that more than three fourths of schools offer mental health, social service, and prevention service options for students and their families.[2] Among psychosocial prevention strategies, there is a broad array of programs available that can be implemented in schools. These include packaged curricula and home-grown programs for use schoolwide and others that target selected children already showing behavior problems or deemed to be at risk for such problems. Each addresses some range of social and emotional factors assumed to cause aggressive behavior or to be instrumental in controlling it (e.g., social skills or emotional self-regulation), and uses one of several broad intervention approaches, with cognitively oriented programs, behavioral programs, social skills training, and counseling/therapy among the most common (see Table 2).

Table 2

Treatment modalities for the four service formats

Modality	Description	Examples	# Universal programs	# Selected programs[a]	# Special programs[a]	# Comp. programs[a]
Behavioral strategies	Techniques, such as rewards, token economies, contingency contracts, and the like to modify or reduce inappropriate behavior.	Good Behavior Game [15]	4	29	13	6
Cognitively-oriented	Focus on changing thinking or cognitive skills; social problem solving; controlling anger, inhibiting hostile attributions, etc.	I Can Problem Solve [16]; Coping Power Program [17]	54	41	17	9
Social skills training	Help youth better understand social behavior and learn appropriate social skills, e.g., communication skills, conflict management, group entry skills, eye contact, "I" statements, etc.	Social skills training [18]; Conflict resolution training [19]	17	26	11	11
Counseling, therapy	Traditional group, individual, or family counseling or therapy techniques.	Mental health intervention [20]; Group counseling [21]	2	26	11	7
Peer mediation	Student conflicts are mediated by a trained st. dent peer	Peer mediation [22]	--	5	--	2

[a]Treatment modalities are not mutually exclusive, except in the universal category where only the focal modality was coded.

In 2003, we published a meta-analysis on the effects of school-based psychosocial interventions for reducing aggressive and disruptive behavior aimed at identifying the characteristics of the most effective programs.[3] That meta-analysis included 172 experimental and quasi-experimental studies of intervention programs, most of which were conducted as research or demonstration projects with significant researcher involvement in program implementation. Although not necessarily representative of routine practice in schools, these programs showed significant potential for reducing aggressive and disruptive behavior, especially for students whose baseline levels were already high. Different intervention approaches appeared equally effective, but significantly larger reductions in aggressive and disruptive behavior were produced by those programs with better implementation, that is, more complete delivery of the intended intervention to the intended recipients.

Since the publication of that review, many new evaluation studies of school-based interventions have become available. The call for schools to implement evidence-based programs has intensified as well. Various resources are available to help schools identify programs with proven effectiveness. Among these resources are the Blueprints for Violence Prevention, the Collaborative for Academic, Social, and Emotional Learning (CASEL), and the National Registry of Evidence-Based Programs and Practices (NREPP) administered by the Substance Abuse and Mental Health Services Administration (SAMSHA). There is, however, little indication that the evidence-based programs promoted to schools through such sources have been widely adopted or that, when adopted, they are implemented with fidelity.[4]

While lists of evidence-based programs can provide useful guidance to schools about interventions likely to be effective in their settings, they are limited by their orientation to distinct program models and the relatively few studies typically available for each such program. A meta-analysis, by contrast, can encompass virtually all credible studies of such interventions and yield evidence about generic intervention approaches as well as distinct program models.

Perhaps most important, it can illuminate the features that characterize the most effective programs and the kinds of students who benefit most. Since many schools already have prevention programs in place, a meta-analysis that identifies characteristics of successful prevention programs can inform schools about ways they might improve those programs or better direct them to the students for whom they are likely to be most effective. Thus, the purpose of the meta-analysis reported here is to update our previous work by adding recent research and further investigate which program and student characteristics are associated with the most effective treatments.

Method

Criteria for Including Studies in the Meta-analysis

Studies were selected for this meta-analysis based on a set of detailed criteria, summarized as follows:

The study was reported in English no earlier than 1950 and involved a school-based program for children attending any grade, pre-Kindergarten through 12th grade.

The study assessed intervention effects on at least one outcome variable that represented either (1) aggressive or violent behavior (e.g., fighting, bullying, person crimes); (2) disruptive behavior (e.g., classroom disruption, conduct disorder, acting out); or (3) both aggressive and disruptive behavior.

The study used an experimental or quasi-experimental design that compared students exposed to one or more identifiable intervention conditions with one or more comparison conditions on at least one qualifying outcome variable.

To qualify as an experimental or quasi-experimental design, a study was required to meet at least one of the following criteria:

Students or classrooms were randomly assigned to conditions;

Students in the intervention and comparison conditions were matched and the matching variables included a pretest for at least one qualifying outcome variable or a close proxy; and

If students or classrooms were not randomly assigned or matched, the study reported both pretest and posttest values on at least one qualifying outcome variable or sufficient demographic information to describe the initial equivalence of the intervention and comparison groups.

Search and Retrieval of Studies

An attempt was made to identify and retrieve the entire population of published and unpublished studies that met the inclusion criteria summarized above. Nearly all of the studies from the original meta-analysis were eligible (pre–post change was also examined in that meta-analysis and some of the studies used for that purpose did not have comparison groups). The primary source of new studies was a comprehensive search of bibliographic databases, including Psychological Abstracts, Dissertation Abstracts International, ERIC (Educational Resources Information Center), United States Government Printing Office publications, National Criminal Justice Reference Service, and MedLine. Second, the bibliographies of recent meta-analyses and literature reviews were reviewed for eligible studies.[5–10] This bibliography was also compared with that from the companion *Guide to Community Preventive Services* (the *Community Guide*) and exchanged citations for studies that they identified and this study did not.[11] Finally, the bibliographies of retrieved studies were themselves examined for candidate studies. Identified studies were retrieved from the library, obtained via interlibrary loan, or requested directly from the author. More than 95% of the reports identified as potentially eligible were obtained and screened through these sources.

Coding of Study Reports

Study findings were coded to represent the mean difference in aggressive behavior between experimental conditions at the posttest measurement. The effect size statistic used for these purposes was the standardized mean difference, defined as the difference between the treatment and control group means on an outcome variable divided by their pooled standard deviation.[12,13] In addition to effect size values, information was coded for each study that described the methods

and procedures, the intervention, and the student samples (coding categories are shown in Table 1). Coding reliability was determined from a sample of approximately 10% of the studies that were randomly selected and recoded by a different coder. For categorical items, intercoder agreement ranged from 73% to 100%. For continuous items, the intercoder correlations ranged from 0.76 to 0.99. A copy of the full coding protocol is available from the first author.

Table 1

Characteristics of the studies with aggressive/disruptive behavior outcomes

Variable	N	%[a]
Subject characteristics		
Gender mix		
All males (> 95%)	43	17
> 60 % males	65	26
50-60% males	89	36
< 50% males	25	10
No males (< 5%)	17	7
Missing	10	4
Age of subjects		
Pre-K & Kindergarten	21	8
6 through 10	106	43
11 through 13	72	29
14 and up	50	20
Predominant ethnicity		
White	77	31
Black	63	25
Hispanic	19	8
Other minority	5	2
Mixed ethnicity	9	4
Missing	72	29
Socioeconomic status		
Mainly low SES	71	29
Working/middle SES	33	13
Mixed, low to middle	28	11
Missing	117	47
Subject risk		
General, low risk	97	39
Selected, risk factors	105	42
Indicated, problem beh.	47	19

[a]Percentages may not add up to 100 because of rounding.

[b]It was often impossible to distinguish between a study with no attrition between pretest and posttest and a study that reported only the number of subjects available at posttest. Thus, although no attrition and unreported attrition are clearly different, they are, of necessity, combined in the same category.

General Analytic Procedures

All effect sizes were multiplied by the small sample correction factor, $1 - (3/4n-9)$, where n is the total sample size for the study, and each was weighted by its inverse variance in all computations.[13,14] The inverse variance weights were computed using the subject-level sample size for each effect size. Because many of the studies used groups (e.g., classrooms, schools) as the unit of assignment to intervention and control conditions, they involved a design effect associated with the clustering of students within classrooms or schools that reduces the effective sample size. The respective study reports provided no basis for estimating those design effects or adjusting the inverse variance weights for them, so they were ignored in the analyses reported here. This should not greatly affect the effect sizes estimates or the magnitude of their relationships to moderator variables, but does assign them somewhat smaller standard error estimates and, hence, larger inverse variance weights than is technically correct. A dummy code identifying these cases was included in the analyses to reveal any differences in findings from these studies relative to those using students as the unit of assignment.

Examination of the effect size distribution identified a small number of outliers with potential to distort the analysis; these were recoded to less extreme values.[13,14] In addition, several studies used unusually large samples. Because the inverse variance weights chiefly reflect sample size, those few studies would dominate any analysis in which they were included. Therefore, the extreme tail of the sample size distribution was recoded to a maximum of 250 students per intervention or control group for the computation of weights. These adjustments allowed us to retain outliers in the analysis, but with less extreme values that would not exercise undue influence on the analysis results.

To create sets of independent effect size estimates for analysis, only one effect size from each subject sample was used in any analysis. When more than one was available, the effect size from the measurement source most frequently represented across all studies (e.g., teachers' reports, self-reports) was selected. The desire was to retain informant as a variable for analysis, so the average across effect sizes from different informants was not used; if there was more than one effect size from the same informant or source, however, their mean value was used.

Finally, many studies provided data sufficient for calculating mean difference effect sizes on the outcome variables at the pretest. In such cases, the posttest effect size was adjusted by subtracting the pretest effect size value. This information was included in the analyses presented below to test whether there were systematic differences between effect sizes adjusted in this way and those that were not.

Analysis of the effect sizes was conducted separately for each program format (described below) and done in several stages. The homogeneity of the effect size distributions using the Q-statistic was tested first.[14] Moderator analyses were then performed to identify the characteristics of the most effective programs using weighted mixed effects multiple regression with the aggressive/disruptive behavior effect size as the dependent variable. In the first stage of this analysis, the influence of study methods on effect sizes was examined. Influential method variables were carried forward as control variables for the next stage of analysis, which examined the relationships between program and student characteristics and effect size. Random effects analysis was used throughout but, in light of the modest number of studies in some categories and the large effect size variance, statistical significance was reported at the alpha=0.10 level as well as the conventional 0.05 level.

Results

Outcomes

The literature search and coding process yielded data from 399 school-based studies. The research studies included in this meta-analysis examined program effects on many different outcomes, ranging from aggression and violence to social skills, academic performance, and self-esteem. Figure 1 presents the mean effect sizes and 95% confidence intervals (CIs) for the most widely represented outcome categories. This report, however, will focus on the outcomes most relevant to school violence prevention, namely aggressive and disruptive behavior.[a]

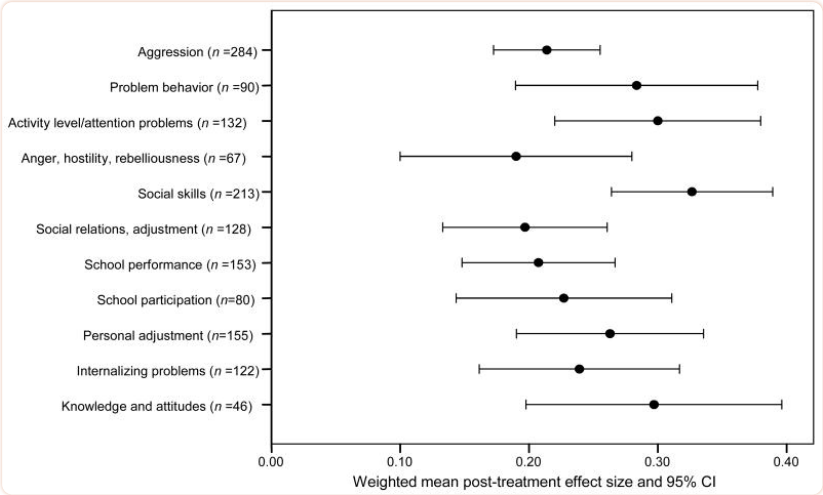

Figure 1

Weighted means and 95% confidence intervals (CI) for the effects of school-based programs on each outcome

The outcome categories shown in Figure 1 are defined as follows. The main outcome of interest is aggressive and disruptive behavior, which involves a variety of negative interpersonal behaviors including fighting, hitting, bullying, verbal conflict, disruptiveness, acting out and the like.[b] The most common type of measure in this category is a teacher-reported survey. Next, there are three categories of behavior problems that are closely related to aggression. These are problem behavior (i.e., measures that include both internalizing and externalizing behaviors like the Child Behavior Checklist [CBCL] Total score; www.aseba.org/products/cbcl6-18.html), activity level/attention problems, and anger/hostility/rebelliousness. Two categories of outcomes relate to social adjustment. The first, and most common after aggression/disruption, includes measures of specific skills, for example, communication skills, social problem solving, conflict resolution skills. Social adjustment, on the other hand, involves measures of how well children get along with their peers, that is, do they have friends, are they well-liked or rejected. The two categories of school outcomes are school performance (e.g., achievement tests, grades) and school participation (e.g., tardiness, truancy, dropout). The personal adjustment category includes measures of self-esteem, self-concept, and other measures of general well-being. Internalizing problems encompasses anxiety, depression, and the like. The final category includes various measures of students' knowledge and attitudes about problem behavior.

As shown in Figure 1, all of these outcomes were positive and statistically significant with mean effect sizes in the 0.20 to 0.35 range. The outcome of primary interest for this meta-analysis, aggressive/disruptive behavior, was most frequently measured via teacher report and showed a mean effect size of 0.21 ($p<0.05$). The results reported in the remainder of this paper pertain only to the effect sizes for these aggressive/disruptive outcomes from the 249 studies that reported them. Our earlier meta-analysis included 172 studies with control group designs and aggressive/disruptive behavior outcomes; thus, the current sample includes an additional 77 studies.

General Study Characteristics

The general characteristics of the 249 studies with aggressive and disruptive behavior outcomes are shown in Table 1. Ninety percent were conducted in the U.S. with nearly 75% done by researchers in psychology or education. Fewer than 20% were conducted prior to 1980 and most were published in peer-reviewed journals (60%), with the remainder reported as dissertations, theses, conference papers, and technical reports.

The student samples reflect the diversity in American schools. Most were comprised of a mix of boys and girls, but there were some all boy samples (17%) and a few all girl samples (7%). Minority children were well represented with over a third of the studies having primarily minority youth; nearly 30%, however, did not report ethnicity information. All school ages were included, from preschool through high school; the average age was around 10. A range of risk levels was also present, from generally low-risk students to those with serious behavior problems. Socioeconomic status was not widely reported, but a range of socioeconomic levels was represented among those studies for which it was reported.

Most studies were conducted as research or demonstration projects with relatively high levels of researcher involvement; however, the number of routine practice programs was increased from eight in the original meta-analysis to 32. Nearly two thirds of the programs were less than 20 weeks in length and about half had service contacts about once per week. Programs were generally manualized and delivered by teachers or the researchers themselves. About 35% of the reports mentioned some difficulties with the implementation of the program. This information, when reported, presented a great variety of relatively idiosyncratic problems, for example, attendance at sessions, dropouts from the program, turnover among delivery personnel, problems scheduling all sessions or delivering them as intended, wide variation between different program settings or providers, results from implementation fidelity measures. This necessitated use of a rather broad coding scheme in which three categories of implementation quality were distinguished: no problems indicated, possible problems (some suggestion of difficulties but little explicit information), and definite problems explicitly reported.

Slightly over 40% of the studies used individual-level random assignment to allocate subjects to treatment and comparison groups. An additional 20% utilized cluster-randomization procedures, usually at the classroom level, although in many cases there were only a few units randomized. The remaining 91 studies used nonrandom procedures to allocate students. Attrition was considerable in some studies, non-existent in others, and averaged about 12%.

Program Format and Treatment Modality

The 249 eligible studies involved a variety of prevention and intervention programs. For purposes of analyzing their effects on student aggressive/disruptive behavior, they were divided into four groups according to their general service format. Programs differ across these groups on a number of methodologic, participant, and intervention characteristics that make it unwise to combine them in a single analysis. The four intervention formats are as follows:

Universal programs. These programs are delivered in classroom settings to all the students in the classroom; that is, the children are not selected individually for treatment but, rather, receive it simply because they are in a program classroom. However, the schools with such programs are often in low socioeconomic status and/or high-crime neighborhoods and, thus, the children in these universal programs may be considered at risk by virtue of their socioeconomic background or neighborhood context.

Selected/Indicated programs. These programs are provided to students who are specifically selected to receive treatment because of conduct problems or some risk factor (typically identified by teachers for social problems or minor classroom disruptiveness). Most of these programs are delivered to the selected children outside of their regular classrooms (either individually or in groups), although some are used in the regular classrooms but targeted on the selected children.

Special schools or classes. These programs involve special schools or classrooms that serve as the usual educational setting for the students involved. Children are placed in these special schools or classrooms because of behavioral or academic difficulties that schools do not want to address in

the context of mainstream classrooms. Included in this category are special education classrooms for behavior disordered children, alternative high schools, and schools-within-schools programs.

Comprehensive/multimodal programs. These programs involve multiple distinct intervention elements (e.g., a social skills program for students and parenting skills training) and/or a mix of different intervention formats. They may also involve programs for parents or capacity building for school administrators and teachers in addition to the programming provided to the students. Within the comprehensive service format, programs were divided into universal and selected/indicated programs. Universal comprehensive programs included multiple treatment modalities, but intervention components were delivered universally to all children in a school or classroom. Selected/indicated comprehensive programs also included multiple modalities, but the children receiving these programs were individually selected for treatment by virtue of behavior problems or risk for such problems. All but one of the programs in this subcategory included services for both students and their parents.

The treatment modalities used in these different service formats varied. However, cognitively oriented approaches and social skills training were common across all four service formats. Cognitively oriented strategies focused on changing thinking patterns, developing social problem solving skills or self control, and managing anger. Social skills training focused on learning constructive behavior for interpersonal interactions, including communication skills and conflict management. Also relatively common among the modalities were behavioral strategies that manipulated rewards and incentives. Counseling for individuals, groups, or families was also represented. Table 2 shows the different treatment modalities used by the programs represented in this meta-analysis and their distribution across the four service formats. For the universal programs, treatment modalities lended themselves to mutually exclusive coding. Treatment modality codes were not mutually exclusive, however, for the selected/indicated, special, and comprehensive service formats. For these service formats, each modality was coded as being present or not present.

Although the universal programs were coded as having a single modality, some did involve multiple treatment components, typically two different types of cognitively oriented programming. Some of the selected/indicated and special programs were coded with more than one treatment component, but were not categorized as comprehensive programs. Unlike the comprehensive programs, they were not billed as comprehensive or multimodal by their authors nor did their multiple components involve different types of treatment and/or different targets (e.g., a school-based cognitive component and a family-based component). The identified multiple treatment components with selected/indicated and special programs were often two types of programming within the same modality (e.g., anger management and social problem solving) or a cognitive component and a social skills component. None of the multiple-component programs in the selected/indicated or special categories involved distinct types of treatment, distinct formats, or multiple targets.

Results for Universal Programs

There were 77 studies of universal programs in the database, all delivered in classroom settings to entire classes of students.[c] Four treatment modalities were represented, as shown in Table 2. Cognitively oriented programs were the primary modality, with some social skills interventions

and a few behavioral and counseling ones. The overall weighted mean effect size on aggressive/disruptive behavior outcomes was 0.21 ($p<0.05$). The test of homogeneity showed significant variability across the effect sizes ($Q_{76}=212$, $p<0.05$).[14] This variation was expected to be associated with the nature of the interventions, students, and methods used in these studies. The first focus was on the relationship between study methods and the intervention effects found by examining the correlation of each method variable with effect size, using random effects inverse variance weights estimated via maximum likelihood.[25]

Table 3 shows the results. Most notable is the lack of significant relationships between the study design variables and effect size. There were only five individual-level random assignment studies of universal programs, so the primary contrast here is between nonrandomized and cluster randomized studies, with neither related to effect size. Only one method variable had a significant correlation—outcome measures reported by the students themselves showed smaller effect sizes than measures from other sources or informants (chiefly teacher reports). Several other variables had modest ($r \geq 0.10$) but nonsignificant correlations with effect size. Outcome measures with more than five items were associated with smaller effect sizes. Effect sizes that were able to be adjusted for pretest differences (by subtracting the pretest effect size) were smaller than unadjusted effect sizes. Greater attrition was also associated with smaller effects. Each of these variables, plus a dummy code for nonrandom assignment, was carried forward to all later analyses to control for the possible influence of method differences on study results.

Table 3

Correlations between study method characteristics and aggressive/disruptive behavior effect sizes for universal programs (N=77)

Method Variable	Correlation
Teacher reported outcome measure	0.07
Self-reported outcome measure	-0.23**
Number of items in outcome measure	-0.19*
Timing of measurement	-0.02
Cluster random assignment	-0.07
Non-random assignment	0.07
Pretest adjustment	-0.13
ES calculated with means/sds (vs. all other methods)	-0.05
Degree of estimation in ES calculation	-0.02
Attrition (% loss)	-0.13
Number of ES aggregated	-0.08

Note: weighted random effects analysis

*$p<0.10$

**$p<0.05$

The next step was to identify student and program characteristics that were associated with effect size while controlling for method variables. To accomplish this, a series of inverse-variance weighted random effects multiple regressions were conducted with each including only a single student or program variable plus the five method variables identified above. These analyses were first run separately in order to identify the relationships between each study characteristic and effect size without the confounding influence of other study characteristics. Table 4 presents the results of these regression analyses.

Table 4

Relationships between Individual study characteristics and aggressive/disruptive behavior effect sizes for universal programs with selected method variables controlled (N=77)

Study characteristic	β(with method controls)[a]
General study characteristics	
Year of publication	-0.03
Unpublished (0) vs. published (1)	0.12
Student characteristics	
Gender mix (% male)	0.07
Age	-0.27**
Mixed or middle SES (0) vs. low SES (1)	0.21*
Researcher role in study	
Routine practice program (1=research, 2=demonstration, 3=routine)	-0.13
Delivery personnel	
Teacher provider	-0.02
Amount & quality of treatment	
Duration of treatment (in weeks; logged)	-0.07
Number of sessions per week (1=less than weekly to 9=daily)	0.09
Implementation problems (1=yes, 2=possible, 3=no)	0.15
Treatment modality	
Cognitively oriented	
Anger management component	0.02
Social problem solving component	0.06
Social skills training	-0.04

Note: weighted random effects analysis; coefficients are standardized.

[a]Method controls: student-reported outcome variable, pretest adjustment, attrition, non-random assignment, number of items in outcome variable.

*p<0.10

**p<0.05

Only two student variables were significantly associated with effect size—age and socioeconomic status. Younger students showed larger effects from universal programming than older students and children with low socioeconomic status showed larger effects than their middle class peers. Several other variables in this analysis had regression coefficients that were modest ($\beta \geq 0.10$) al-

though nonsignificant. Published studies, research and demonstration programs (versus routine practice), and well-implemented programs all showed somewhat larger effect sizes than studies without these characteristics.

Note that Table 4 reports the relationship between effect size and each of the three most common treatment modalities for universal programs. The cognitively oriented programs were separated into two groups: anger management programs and social problem solving programs. These were the most frequent types of cognitively oriented programs and were not mutually exclusive. The third category included the social skills programs. None of these treatment modalities was associated with significantly larger or smaller effect sizes relative to the others.

To examine the independent influence of all the variables identified so far as potential moderators of intervention effects, the significant variables from Table 4, as well as those with individual regression coefficients larger than 0.10 and the five method controls, were carried forward into a summary regression analysis. As shown in Table 5, only student socioeconomic status was significant in this model, although several other variables showed nonsignificant regression coefficients of ≥0.10. As in the individual variable analysis above, students with low socioeconomic status achieved significantly greater reductions in aggressive and disruptive behavior from universal programs than middle class students. In addition, published studies, younger students, research and demonstration programs, and implementation quality were all modestly associated with larger effect sizes, although these relationships did not reach statistical significance.

Table 5

Regression model for effect size moderators for universal programs (N=7)

Study characteristic	β
Method Characteristics	
Self-reported dependent measure	-0.13
Pretest adjustment	0.05
Attrition	0.04
Non-random assignment	0.06
Number of items in outcome measure	-0.18
General study characteristics	
Unpublished (0) vs. published (1)	0.19
Student characteristics	
Age	-0.18
Mixed or middle SES (0) vs. low (1)	0.27**
Researcher role in study	
Routine practice program (1=research, 2=demo, 3=routine)	-0.10
Amount andality of treatment	
Implementation quality	0.14

Note: weighted random effects analysis; coefficients are standardized.

*p<0.10

**p<0.05

Results for Selected/Indicated Programs

There were 108 studies of selected/indicated programs that targeted interventions to individually identified children. Nearly all of these programs were "pull-out" programs delivered outside the classroom to small groups or individual students. The overall random effects mean effect size for these programs was 0.29 (p <0.05). Five treatment modalities were identified among these programs, as described in Table 2. As with the universal programs, the most common programs were cognitively oriented, although behavioral strategies, social skills training, and counseling programs were well represented. Many of the behavioral programs for selected students involved an in-class component (e.g., behavioral contracts monitored by the teacher).

The homogeneity test of the effect sizes showed significant variability across studies (Q_{108}=300, p<0.05) and the analysis of the relationships between effect size and methodologic and substantive characteristics of the studies proceeded much the same as for the universal programs. First,

the correlation of each method variable with the aggressive/disruptive behavior effect sizes was examined (Table 6). Here also the study design was not associated with effect size—random assignment studies did not show appreciably smaller or larger effects than nonrandomized studies. Note that for the selected/indicated programs the design contrast was primarily between individual-level randomization and nonrandomization; there were only six cluster randomized studies. The two method variables that did show significant zero-order relationships with effect size were outcome measures with more than five items and attrition, both associated with smaller effect sizes. Adjustment of effect sizes for pretest differences was the only other method variable with a correlation larger than 0.10 with effect size, but it did not reach statistical significance. Four method variables were carried forward into additional analyses: random assignment, pretest adjustment, number of items in the outcome measure, and attrition.

Table 6

Correlations between study method characteristics and aggressive/disruptive behavior effect sizes for selected/indicated pull-out programs (N=108)

Method variable	Correlation
Teacher reported outcome measure	-0.00
Archival outcome measure	0.06
Observational outcome measure	-0.00
Number of items in outcome measure	-0.19*
Timing of measurement	0.07
Random assignment	-0.01
Pretest adjustment	-0.11
ES calculated with means/sds (vs. all other methods)	-0.09
Degree of estimation in ES calculation	-0.09
Attrition (% loss)	-0.22**
Number of ES aggregated	0.05

Note: weighted random effects analysis

*p<0.10

**p<0.05

Table 7 shows the regression coefficients from a series of regression analyses, each of which included the four method control variables and a single substantive variable. Five student and program variables had significant relationships with effect size in these analyses. Higher-risk subjects showed larger effect sizes than lower risk subjects, though, with the selected/indicated programs, very few low-risk children were involved. The distinction here is mainly between indicated students who are already exhibiting behavior problems and selected students who have risk factors

that may lead to later problems. Regarding the intervention programs, individual treatment (versus group) and programs with higher quality implementation were associated with larger effects. In addition, programs using behavioral strategies produced significantly greater reductions in aggressive/disruptive behavior than the other modalities.

Table 7

Relationships between individual study characteristics and aggressive/disruptive behavior effect sizes for selected/indicated pull-out programs with method variables controlled (N=108)

Study characteristic	β(with method controls)[a]
General study characteristics	
Year of publication	-0.12
Unpublished (0) vs. published (1)	-0.16
Student characteristics	
Gender mix (% male)	0.05
Age	0.04
Mixed or middle SES (0) vs. low (1)	0.05
Risk level	0.23**
Researcher role in study	
Routine practice program (1=research, 2=demonstration, 3=routine)	0.09
Delivery personnel	
Researcher provider	0.05
Teacher provider	0.01
Service professional provider	0.03
Amount, format, & quality of treatment	
Manualized (1) vs. unstructured treatment (2)	0.09
Group treatment	-0.16*
Individual treatment	0.17*
Duration of treatment (in weeks; logged)	0.07
Number of sessions per week (1=less than weekly to 7=daily)	0.02
Implementation problems (1=yes, 2=possible, 3=no)	0.15*
Treatment modality	
Cognitively oriented	
Anger management component	-0.04
Social problem solving component	-0.07
Social skills training	-0.06
Counseling	-0.02
Behavioral strategies	0.20**

Note: weighted random effects analysis; coefficients are standardized.

[a]Method controls: pretest adjustment, attrition, random assignment, number of items in the outcome measure.

*$p<0.10$

**p<0.05

The four significant student and program variables and the two with individual regression coefficients greater than 0.10, along with the four method control variables, were included in the final summary regression model shown in Table 8. Two methodologic characteristics were significantly associated with smaller effects—greater attrition and outcome variables with more than five items. The risk variable was also significant; programs achieved larger effects with higher risk students. Socioeconomic status, although not related to effect size, was significantly correlated with risk such that higher risk students tended to be of lower socioeconomic status. Individual treatments were no longer significantly different from other forms of delivery, although the relationship still favored individual treatments. Better implemented programs produced significantly larger effects than poorly implemented ones. Finally, programs using behavioral strategies were more effective than those which used other modalities.

Table 8

regression model for effect size moderators for selected/indicated pull-out programs (N=108)

Study characteristic	β
Method characteristics	
Random assignment	.05
Pretest adjustment	-.03
Attrition (% loss)	-.21**
Number of items in outcome measure	-.15*
General study characteristics	
Year of publication	-.10
Published (1) vs. unpublished (0)	-.08
Student characteristics	
Risk level (1=general; 2=at-risk; 3=indicated)	.19**
Amount, format, & quality of treatment	
Individual treatment	.11
Implementation problems (1=yes, 2=possible, 3=no)	.18**
Treatment modality	
Behavioral strategies	.15*

Note: weighted random effects analysis; coefficients are standardized.

*p<.10

**p<.05

Results for Special Schools or Classes

There were 43 studies of programs delivered in special schools or classrooms. These programs generally involved an academic curriculum plus programming for social or aggressive behavior. The students typically had behavioral (and often academic) difficulties that resulted in their placement outside of mainstream classrooms. The mean aggressive/disruptive behavior effect size for these programs was 0.11 ($p<0.10$). The Q test was significant ($Q_{42}=82$, $p<0.05$), indicating that the distribution of effect sizes was heterogeneous. About 40% of the studies of special programs assigned students to intervention and control conditions at the classroom level, while the remaining 60% used individual-level assignment. As a result, there may be a design effect associated with the clustering of students within classrooms that overstates the significance, although the overall effect size and the regression coefficients presented below should not be greatly affected.

The correlations between the method variables and effect sizes are shown in Table 9. Effect sizes adjusted for pretest differences were significantly larger than effect sizes that were not adjusted, a contrast with the universal and selected/indicated programs where pretest adjustments were associated with smaller effect sizes. Although not significant, studies with individual-level random assignment were associated with smaller effects than studies that used other assignment methods and greater attrition was associated with smaller effect sizes. In addition, self-reported outcomes tended to produce smaller effect sizes.

Table 9

correlations between study method characteristics and aggressive/disruptive behavior effect sizes for special programs (N=43)

Method variable	Correlation
Teacher reported outcome measure	.01
Self reported outcome measure	-.24
Number of items in outcome measure	.04
Random assignment	-.20
Cluster random assignment	.15
Non-random assignment	.08
Pretest adjustment	.30**
ES calculated with means/sds (vs. all other methods)	-.03
Attrition (% loss)	-.24
Number of ES aggregated	-.05

Note: weighted random effects analysis

*p<.10

**p<.05

For the next stage of analysis, the self reported outcome, pretest adjustment, random assignment, and attrition variables were carried forward as method controls in regression analyses with individual study characteristics (Table 10). Two variables were significant, whether an in-class or pull-out program and implementation quality. In one form of the special programs, students were assigned to special education classes or schools and the program was delivered entirely in the classroom setting. The other form involved students in special education classrooms who were pulled out of class for additional small group treatments. The programs delivered in classroom settings showed larger reductions in aggressive/disruptive behavior than the pull-out programs. Also, as in other analyses, better implemented programs showed larger effects. Two treatment modalities were tested in this model, cognitively oriented strategies and schools-within-schools programs, and neither were found to be significant. The cognitively oriented programs were generally similar to cognitive programs delivered within the universal and selected/indicated service formats. Schools within schools were generally delivered with middle and high schools students and consisted of groups of students who were placed together for most or all of their instruction. Schools within schools are often housed in a separate building or set of classrooms on a larger campus, and are characterized by smaller student–teacher ratios and more individualized attention. In many cases, the schools within schools programs included here were designed for behavior problem youth.[26, 27]

Table 10

Relationships between individual study characteristics and aggressive/disruptive behavior effect sizes with method variables controlled for special programs (N=43)

Study characteristic	β(with method controls)[a]
General study characteristics	
Year of publication	-.22
Unpublished (0) vs. published (1)	.04
Student characteristics	
Gender mix (% male)	.07
Age	-.03
Mixed or middle SES (0) vs. low (1)	-.08
Risk level	.25
Researcher role in study	
Routine practice program (1=research, 2=demo, 3=routine)	.01
Delivery personnel	
Teacher provider	.05
Amount, format, & quality of treatment	
Manualized (1) vs. unstructured treatment (2)	-.14
In-class (1) vs. pull-out treatment (2)	-.38**
Duration of treatment (in weeks; logged)	-.06
Number of sessions per week (0=less than daily, 1=daily)	.17
Implementation problems (0=yes, 1=no)	.42**
Treatment modality	
Cognitively-oriented	-.08
Schools within schools component	.02

Note: weighted random effects analysis; coefficients are standardized.

[a]Method controls: self-reported outcome measure, pretest adjustment, attrition, non-random assignment.

*p<.10

**p<.05

Results for Comprehensive or Multimodal Programs

There were only 21 studies of comprehensive programs in this database, distinguished by their multiple treatment components and formats. The average number of distinct treatment components for these programs was four, whereas the universal and selected/indicated programs typi-

cally had one treatment component.[d] The studies of comprehensive programs tended to involve larger samples of students than the other program formats and, like the special and universal programs, a larger proportion of cluster randomizations. Thus, the significance of the mean reported below is overstated. Comprehensive programs were generally longer than the universal and selected/indicated programs. The modal program covered an entire school year and almost half of the programs were longer than 1 year. In contrast, the average program length for universal and selected/indicated programs was about 20 weeks.

The overall mean effect size for the comprehensive programs was .05 and was not statistically significant. Students who participated in comprehensive programs were no better off than students who did not. In addition, the Q-test test showed that the distribution of effect sizes was homogeneous (Q_{20}=28, p >0.10). However, the Q-test has relatively low statistical power with small numbers of studies so, despite the nonsignificant effect size heterogeneity, the correlations between study method and substantive characteristics and effect size were examined. Table 12 shows significant bivariate relationships for nonrandomized assignment (larger effect sizes) and cluster randomization (smaller effect sizes). Among the program variables, longer treatments and more frequent sessions per week were associated with larger effect sizes. Universally delivered programs showed larger effects than pull-out programs. Table 13 shows that when the variables with significant correlations with effect size were included together in a regression model, only universally delivered (versus pull-out) programs and frequency of sessions per week showed significant independent relationships to effect size. Recall that the comprehensive programs were divided into those that were universally delivered to all students regardless of risk (n=12) and those that involved students individually selected for problem behavior or risk for such behavior (n=9). Although the mean effect size for all comprehensive programs was small and nonsignificant, universally delivered programs and those with more frequent treatment contacts tended to produce larger reductions in aggressive and disruptive behavior.

Table 12

Correlations between study characteristics and aggressive/disruptive behavior effect sizes for comprehensive programs (N=21)

Study Variable	Correlation
Teacher reported outcome variable	.07
Number of items in outcome measure	.27
Number of ES aggregated	-.04
Random assignment	-.05
Non-random assignment	.42**
Cluster random assignment	-.33
Attrition (% loss)	.25
Publication year	-.10
Published (1) vs. unpublished (0)	-.23
Role of evaluator (1=delivered tx; 4=research role only)	-.14
Treatment duration (weeks)	.34*
Frequency of sessions per week	.44**
Implementation quality	.17
Universal (1) vs. pull-out (2) format	-.34
Low SES (vs. mixed or middle class)	-.08
Risk level of subjects (low to high)	-.11
Age	.10
Gender mix (% male)	-.12

Note: weighted random effects analysis

*p<.10

**p<.05

Table 13

regression model for effect size moderators for comprehensive programs (N=21)

Study Characteristic	β
Method characteristics	
Non-random assignment	-.03
Format of program	
Universal (1) vs. Pull-out (2)	-.43*
Amount of treatment	
Frequency of sessions per week	.53*
Program duration (weeks)	-.02

Note: weighted random effects analysis; coefficients are standardized.

**p<.05

*p<.10

Summary and Conclusion

The issue addressed in this paper is the effectiveness of programs for preventing or reducing such aggressive and disruptive behaviors as fighting, bullying, name-calling, intimidation, acting out, and disruptive behaviors occurring in school settings. The main finding is that, overall, the school-based programs that have been studied by researchers (and often developed and implemented by them as well) generally have positive effects for this purpose. The most common and most effective approaches are universal programs delivered to all the students in a classroom or school and targeted programs for selected/indicated children who participate in programs outside of their regular classrooms. The universal programs that were included in the analysis mainly used cognitive approaches, so it is not clear whether their generally positive effects stem more from the universal service format or the cognitively oriented treatment modality. Cognitively oriented approaches were also the most frequent among the selected/indicated programs, but many did use behavioral, social skills, or counseling treatment modalities. Other than somewhat larger effects for programs with a behavioral component, differential use of these modalities was not associated with differential effects. This suggests that it may be the selected/indicated program format that is most important but does not rule out the possibility that the small number of treatment modalities used with that format are especially effective ones.

The mean effect sizes of 0.21 and 0.29 for universal and selected/indicated programs, respectively, represent a decrease in aggressive/disruptive behavior that is not only statistically significant but likely to be of practical significance to schools as well. Suppose, for example, that approximately 20% of students are involved in some version of such behavior during a typical school year. This is a plausible assumption according to the *Indicators of School Crime and Safety: 2005*, which re-

ports that 13% of students aged 12–18 were in a fight on school property, 12% had been the targets of hate-related words, and 7% had been bullied.[28] Effect sizes of 0.21 and 0.29 represent reductions from a base rate prevalence of 20% to about 15% and 13%, respectively, that is, 25%–33% reductions. The programs of above average effectiveness, of course, produce even larger decreases.

The substantial similarity of the mean effect sizes across service formats and treatment modalities for the universal and selected/indicated programs suggests that schools may choose from a range of such programs with some confidence that whatever they pick will be about as effective as any other choice. In the absence of evidence that one modality is significantly more effective at reducing aggressive and disruptive behavior than another, schools might benefit most by considering ease of implementation when selecting programs and focusing on implementation quality once programs are in place. The coding of implementation quality, albeit crude, was associated with larger effect sizes for all four treatment formats, although statistically significant only for selected/indicated and special programs. A very high proportion of the studies in this meta-analysis, however, were research or demonstration projects in which the researchers had a relatively large direct influence on the service delivery. Schools adopting these programs without such engagement may have difficulty attaining comparable program fidelity, a concern reinforced by evidence of frequent weak implementation in actual practice[4]. The best choice of a universal or selected/indicated program for a school, therefore, may be the one they are most confident they can implement well.

Another significant factor that cut across the universal and selected/indicated programs was the risk level of the students receiving the intervention. Larger treatment effects were achieved with higher risk students. For the universal programs, the greatest benefits appeared for students from economically disadvantaged backgrounds while, for the selected/indicated programs, it was students already exhibiting problematic behavior that showed the largest effects. Universal programs did not specifically select students with individual risk factors or behavior problems, although many students were of low socioeconomic status and there were most likely some behavior problem students in the classrooms that received universal interventions. And, although socioeconomic status was not significant in the analysis of selected/indicated programs, the weighted correlation between risk and socioeconomic status for the selected/indicated students was significant.[41] These findings reinforce the truism that a program cannot have large effects unless there is sufficient problem behavior, or risk for such behavior, to allow for significant improvement.

The programs in the category that are called comprehensive, in contrast to the universal and selected/indicated programs, were surprisingly ineffective. On the face of it, combinations of universal and pull-out treatment elements and multiple intervention strategies would be expected to be at least as effective, if not more so, than less multifaceted programs. Their small and nonsignificant mean effect size raises questions about the value of such programs. It should be noted, however, that most of these were long-term schoolwide programs. It may be that this broad scope is associated with some dilution of the intensity and focus of the programs so that students have less engagement with them than with the programs in the universal and selected/indicated categories.

It may also be relevant that proportionately fewer of the programs in this category involved the cognitively oriented treatment modalities that were the most widely represented ones among the universal and selected/indicated programs. This is an area that clearly warrants further study.

The most distinctive programs in this collection were those for students in special education and other such atypical school settings. The mean effect size for these programs was modest though statistically significant. These results also are somewhat anomalous. One of the signal characteristics of students in these settings is a relatively high level of behavior problems or risk for such problems, thus there should be ample room for improvement. On the other hand, the special school settings in which they are placed can be expected to already have some programming in place to deal with such problems. The control conditions in these studies would thus reflect the effects of that practice-as-usual situation with less value added provided by additional programming of the sort examined in these studies. Alternatively, however, the add-on programs studied in these cases may have been weaker than those found in the selected/indicated format or the more serious behavior problems of students in these settings may be more resistant to change. Here too the issues that warrant further study.

A particular concern of our earlier meta-analysis was the smaller effects of routine practice programs in comparison to those of the more heavily represented research and demonstration programs.[3] Routine practice programs are those implemented in a school on an ongoing routine basis and evaluated by a researcher with no direct role in developing or implementing the program. Research and demonstration programs are mounted by a researcher for research or demonstration purposes with the researcher often being the program developer and heavily involved in the implementation of the program, although somewhat less so for demonstration programs. In the present meta-analysis, somewhat more studies of routine programs were included and it is reassuring that their mean effect sizes, although smaller than those for research and demonstration programs, were not significantly smaller. As shown in Tables 4, 7, and 10, routine practice programs did not show significantly better or worse outcomes than research and demonstration programs for universal, selected/indicated or special programs.[d] Only 32 of the 249 studies in this meta-analysis examined routine practice programs, however, with 13 in the universal format, 11 in the selected/indicated pullout format, and 7 in the special format. This number dramatizes how little evidence exists about the actual effectiveness, in everyday real-world practice, of the kinds of school-based programs for aggressive/disruptive behavior represented in this review.

Table 11

Regression model for effect size moderators for special programs (N=43)

Study Characteristic	β
Method characteristics	
Self reported outcome measure	.18
Random assignment	.02
Pretest adjustment	.28
Attrition (% loss)	-.27
General study characteristics	
Year of publication	-.04
Student characteristics	
Risk level	.21
Amount, format, & quality of treatment	
In-class (1) vs. pull-out treatment (2)	-.24
Implementation problems (0=yes, 1=no)	.32**

Note: weighted random effects analysis; coefficients are standardized.

*p<.10

**p<.05

Acknowledgments

The research reported in this article was supported by grants from the National Institute of Mental Health, the National Institute of Justice, and the Centers for Disease Control and Prevention. The authors gratefully acknowledge the assistance of Kerstin Blomquist, Rose Vick, Dina Capitani, Gabrielle Chapman, Carole Cunningham, Jorie Henrickson, Young-il Kim, Jan Morrison, Kushal Patel, Shanna Ray, and Judy Formosa.

Footnotes

A bibliography of studies included in the meta-analysis is available from the first author or on the following website: www.vanderbilt.edu/CERM.

[a]Studies otherwise eligible but without aggressive/disruptive behavior outcomes were coded as part of a larger project. Thus, 399 studies appear in Figure 1, while only 249 are represented in the primary analysis of aggressive and disruptive behavior effect sizes.

[b]Ideally, we would have liked to examine program effects only on aggressive behavior. However, almost none of the measures that call themselves aggressive behavior measures focus solely on physically aggressive interpersonal behavior. Many include disruptiveness, acting out, and other forms of behavior problems that are negative, but not necessarily aggressive.

[c]There were three universal programs that were delivered to entire classrooms, but certain children (those at risk) were selected for analysis. These were retained in the universal format category because the experiences of these children were more similar to the universal programs than the selected/indicated programs.

[d]There was only one routine practice program with a comprehensive format; thus, the routine practice variable was not included in any analyses of comprehensive programs.

No financial conflict of interest was reported by the authors of this paper.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

References

1. Gottfredson GD, Gottfredson DC, Czeh ER, Cantor D, Crosse S, Hantman I. *National study of delinquency prevention in schools.* Ellicott City (MD): Gottfredson Associates, Inc.; 2000. Final Report, Grant No. 96-MU-MU-0008. Available from: www.gottfredson.com. [Google Scholar]

2. Brener ND, Martindale J, Weist MD. Mental health and social services: results from the School Health Policies and Program Study. *J of Sch Health.* 2001;71:305–12. [PubMed] [Google Scholar]

3. Wilson SJ, Lipsey MW, Derzon JH. The effects of school-based intervention programs on aggressive and disruptive behavior: a meta-analysis. *J Consult Clin Psychol.* 2003;71:136–49. [PubMed] [Google Scholar]

4. Gottfredson DC, Gottfredson GD. Quality of school-based prevention programs: results from a national survey. *J Res Crime Delinq.* 2002;39:3–35. [Google Scholar]

5. Durlak JA. *School-based prevention programs for children and adolescents.* Thousand Oaks: Sage; 1995. [Google Scholar]

6. Durlak JA. Primary prevention programs in schools. *Adv Clin Child Psych.* 1997;19:283–318. [Google Scholar]

7. Gansle KA. The effectiveness of school-based anger interventions and programs. *J Sch Psychol.* 2005;43:321–341. [Google Scholar]

8. Lösel F, Beelmann A. Effects of child skill training in preventing antisocial behavior: a systematic review of randomized evaluations. *Ann Am Acad Pol Soc Sci.* 2003;587:84–109. [Google Scholar]

9. Sukhodolsky DG, Kassinove H, Gorman BS. Cognitive–behavioral therapy for anger in children and adolescents: a meta-analysis. *Aggress Violent Beh.* 2004;9:247–269. [Google Scholar]

10. Wilson DB, Gottfredson DC, Najaka SS. School-based prevention of problem behaviors: a meta-analysis. *J of Quantitative Criminology.* 2001;17:247–272. [Google Scholar]

11. Hahn R, Fuqua-Whitley D, Wethington H, et al. Effectiveness of universal school-based programs to prevent violence: a systematic review. *Am J Prev Med.* 2006 Under review. [Google Scholar]

12. Cohen J. *Statistical power analysis for the behavioral sciences.* 2nd. Hillsdale: Lawrence Erlbaum; 1988. [Google Scholar]

13. Lipsey MW, Wilson DB. *Practical meta-analysis.* Thousand Oaks: Sage; 2001. [Google Scholar]

14. Hedges LV, Olkin D. *Statistical methods for meta-analyses.* San Diego: Academic Press; 1985. [Google Scholar]

15. Dolan LJ, Kellam SG, Brown CH, Werthamer-Larsson L, Rebok GW, Mayer LS, Laudolff J, Turkkan JS, Ford C, Wheller L. The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *J Appl Dev Psychol.* 1993;14:317–345. [Google Scholar]

16. Shure MB, Spivack G. Interpersonal problem solving as a mediator of behavioral adjustment in preschool and kindergarten children. *J Appl Dev Psychol.* 1980;1:29–44. [Google Scholar]

17. Lochman JE, Wells KC. The Coping Power program at the middle-school transition: universal and indicated prevention effects. *Psychol Addict Behav.* 2002;16 4:S40–S54. [PubMed] [Google Scholar]

18. Gresham FM, Nagle RJ. Social skills training with children: responsiveness to modeling and coaching as a function of peer orientation. *J Consult Clin Psych.* 1980;48:718–729. [PubMed] [Google Scholar]

19. Roseberry LL. *An applied experimental evaluation of conflict resolution curriculum and social skills development [dissertation]* Chicago (IL): Loyola Univ.; 1997. [Google Scholar]

20. Nafpaktitis M, Perlmuter BF. School-based early mental health intervention with at-risk students. *School Psychol Rev.* 1998;27:420–432. [Google Scholar]

21. Bestland DE. *A controlled experiment utilizing group counseling in four secondary schools in the Milwaukee public schools [dissertation]* Milwaukee (WI): Marquette Univ.; 1968. [Google Scholar]

22. Miller PH. *The relative effectiveness of peer mediation: children helping each other to solve conflicts [dissertation]* Oxford (MI): Univ. of Mississippi; 1995. [Google Scholar]

23. Catalano RF, Mazza JJ, Harachi TW, Abbott RD, Haggerty KP, Fleming CB. Raising healthy children through enhancing social development in elementary school: results after 1.5 years. *J School Psychol.* 2003;41:143–164. [Google Scholar]

24. Conduct Problems Prevention Research Group. Initial impact of the Fast Track prevention trial for conduct problems: I. the high-risk sample. *J Consult Clin Psych.* 1999;67:631–647. [PMC free article] [PubMed] [Google Scholar]

25. Raudenbush SW. Random effects models. In: Cooper H, Hedges LV, editors. *The handbook of research synthesis.* New York: Russell Sage Foundation; 1994. pp. 301–21. [Google Scholar]

26. Weinstein RS, Soule CR, Collins R, Cone J, Mehlhorn M, Simontacchi K. Expectations and high school change: teacher–researcher collaboration to prevent school failure. *Am J Commun Psychol.* 1991;19:333–363. [PubMed] [Google Scholar]

27. Hudley CA. Assessing the impact of separate schooling for African-American male adolescents. *J Early Adolescence.* 1995;15:38–57. [Google Scholar]

28. DeVoe JF, Peter K, Kaufman P, et al. *Indicators of school crime and safety: 2004.* Washington: U S Departments of Education and Justice; 2004. NCES 2005-002/NCJ 205290. [Google Scholar]