The Effectiveness of Whole-School Antibullying Programs: A Synthesis of Evaluation Research

J. David Smith and Barry H. Schneider
University of Ottawa

Peter K. Smith and Katerina Ananiadou
University of London

Abstract. Bullying is a serious problem in schools, and school authorities need effective solutions to resolve this problem. There is growing interest in the whole-school approach to bullying. Whole-school programs have multiple components that operate simultaneously at different levels in the school community. This article synthesizes the existing evaluation research on whole-school programs to determine the overall effectiveness of this approach. The majority of programs evaluated to date have yielded nonsignificant outcomes on measures of self-reported victimization and bullying, and only a small number have yielded positive outcomes. On the whole, programs in which implementation was systematically monitored tended to be more effective than programs without any monitoring.

Bullying is a particularly vicious kind of aggressive behavior distinguished by repeated acts against weaker victims who cannot easily defend themselves (Farrington, 1993; Smith & Brain, 2000). Its consequences are severe, especially for those victimized over long periods of time. Bullying is a complex psychosocial problem influenced by a myriad of variables. The repetition and imbalance of power involved may be due to physical strength, numbers, or psychological factors.

Both bullies and victims evidence poorer psychological adjustment than individuals not involved in bullying (Kumpulainen, Raesaenen, & Henttonen, 1999; Nansel et al., 2001). Children who bully tend to be involved in alcohol consumption and smoking, have poorer academic records than noninvolved students, display a strong need for dominance, and

show little empathy for their victims (Roberts & Morotti, 2000). Bullying may be a means of increasing one's social status and access to valued resources, such as the attention of opposite-sex peers (Pellegrini, 2001). Victims tend to be socially isolated, lack social skills, and have more anxiety and lower self-esteem than students in general (Olweus, 1997). They also tend to have a higher than normal risk for depression and suicide (e.g., Sourander, Helstelae, Helenius, & Piha, 2000). A subgroup of victims reacts aggressively to abuse and has a distinct pattern of psychosocial maladjustment encompassing both the antisocial behavior of bullies and the social and emotional difficulties of victims (Glover, Gough, Johnson, & Cartwright, 2000). Bullying is a relatively stable and long-term problem for those involved, particularly children fitting the profile

Address correspondence concerning this article to Dr. David Smith, Faculty of Education, University of Ottawa, P.O. Box 450, Stn. A, Ottawa, Ontario K1N 6N5; E-mail: jdsmith@uottawa.ca

of a bully-victim (Olweus, 1994). Problems with abusive and antisocial behavior during childhood often extend into adulthood (Farrington & Hawkins, 1991).

Parents and home environments also can promote bullying (Smith & Myron-Wilson, 1998). Children who bully tend to come from homes where aggression is a favored problem-solving method, negative emotional attitudes (e.g., lack of warmth and involvement) are common, and the children are encouraged to fight back when harassed (Glover et al., 2000; Roberts & Morotti, 2000). Conversely, children can be protected against serious problems associated with bullying by authoritative parents who communicate love and warmth, set appropriate limits, and use nonphysical punishment to correct misbehavior (e.g., Olweus, 1993).

Bullying is a group process that involves and is enabled by many players in addition to the individual bullies and victims. Bullying usually occurs in the presence of peers, who can adopt a variety of roles, such as remaining neutral during a bullying incident, assisting and encouraging the bully, or aiding or consoling the victim. Unfortunately, this last role is rarely adopted by children. The actions of peers in the vicinity of bullying incidents typically support the bullying behavior rather than stop it (Salmivalli, 1999; Sutton & Smith, 1999). It has been argued within a social learning perspective that bullies are reinforced for their actions by the attention and encouragement they receive from the onlookers and that peers are more likely to imitate rather than censure bullies' behavior because they see the rewards that accrue to the bullies (Craig, Pepler, & Atlas, 2000). The social processes underlying bullying are not, however, the exclusive domain of children. Adults in the immediate environment (e.g., teachers) can have a direct effect on the bullying process. For example, they may become vigilant and intervene when appropriate or, alternatively, overlook or ignore bullying when it occurs. The group process associated with bullying bolsters the argument for a systemic approach to its correction.

The multiple causes of bullying suggest multiple avenues for possible intervention. One avenue of intervention is designated as whole-school. The whole-school approach is predicated on the assumption that bullying is a systemic problem, and, by implication, an intervention must be directed at the entire school context rather than just at individual bullies and victims. One advantage of the whole-school approach is that it avoids the potentially problematic stigmatization of either bullies or victims. It also circumvents the potential for cross-fertilization of beliefs that aggression is legitimate among aggressive children brought together for intervention in some forms of group counseling and social skills training. This negative side effect has occurred in some programs intended for adolescents in the U.S. (Dishion, McCord, & Poulin, 1999).

The Olweus Bullying Prevention Program (Olweus, 1993) was the first comprehensive whole-school intervention implemented on a large scale and systematically evaluated. The interventions examined in our review share the core features of the original Olweus program. Within this approach, it is considered essential that all members of the school community, including school staff, pupils, and parents, be sensitized with basic information about what bullying is and how they should respond to it. A clear, consistent policy involving nonphysical consequences for bullying needs to be developed and implemented. This entails communication among the adults in the victims' lives and among the adults in the school, who must supervise children's interactions actively and react as a team. Curricular activities are designed to instill antibullying attitudes in all children and assist them in developing prosocial conflict resolution skills. Finally, individualized interventions are developed for children directly involved in bullying as either victims or bullies. Whole-school interventions often have not been applied to more general aspects of school climate, such as general quality of communication, interpersonal relationships, organization, and academic emphasis. Ortega and Lera (2000) have suggested that these ecological factors influence bullying, but they are not usually incorporated into systematic intervention packages that have been subjected to empirical scrutiny.

The purpose of the present research synthesis is to assess the evidence for the efficacy

of whole-school antibullying programs. This study began with the most basic question that probably occurs to practitioners, namely whether the whole-school approach leads to a reduction in bullying and victimization. To the limited extent permitted by the original research reports, we investigated whether effectiveness depended on the age of the students, the characteristics of the intervention, or characteristics of the research design. In answering these broad questions, we averaged across studies that share the general whole-school philosophy, but that differ in some features that may be important in guiding practice. The option of basing the review only on studies providing the best empirical evidence was rejected because this would have resulted in the elimination of most of the database. Unfortunately, random assignment to treatment and appropriate control conditions and multimethod outcome assessment using measures with proven validity and reliability are the exception rather than the rule. This article concludes by considering these limitations in more detail and discussing their implications for practice and future research.

Method

Document Search Strategy

Criteria for inclusion in this review were: (a) The study pertained to a systematic evaluation of a whole-school antibullying intervention; (b) the report provided quantitative outcome data on victimization and/or bullying in schools; and (c) the study was conducted in more than one classroom. These research reports were located by searching relevant databases (PsychINFO, ERIC, and Dissertation Abstracts). Unfortunately, the term "bullying" is not in the thesaurus of these databases. Although "victimization" is a recognized subject term, it is not specific to victims of bullying. Therefore, the search terms "bullying and intervention" and "anti-bullying" were used to perform a keyword search. The first term yielded a total of 241 documents across the three databases, and the term "anti-bullying" yielded a total of 82 documents. All documents identified in this search were individually examined to determine whether or not they met

inclusion criteria for the study. Documents identified in this part of the search complemented several unpublished research reports already located. In addition to this strategy, an attempt was made to identify all researchers in North America and Europe who have conducted research on bullying generally and/or antibullying interventions in particular. These researchers were identified by consulting recent conference programs of the International Society for Research on Aggression and recent volumes of the journal *Aggressive Behavior*. Finally, the reference lists of all relevant documents at hand were scanned to find researchers not identified through other search strategies. The search continued until the point of saturation was reached and no new names were appearing. All researchers identified through this process were contacted directly and asked to furnish a copy of any research report, published or unpublished, they had authored that evaluated an intervention program intended to reduce the incidence of victimization and/or bullying in schools. This search for pertinent documents ended in December 2002 at which point 14 studies had been located that met inclusion criteria.

Coding of Study Features

Each study was coded for relevant methodological and program features designated by the authors. These included (a) the components of the interventions (e.g., school policy, teacher workshops), (b) the design of the studies (whether they were controlled or uncontrolled, and randomness of group assignment), (c) the nature of outcome data (e.g., observations, self-reports of bullying), (d) participants' ages and grades, and (e) the time intervals for data collection. Two independent research assistants coded each document included in this study (with the exception of two studies not written in English). Interrater agreement for coding of methodological features and study outcomes was 94% and for intervention components was 85%.

Results

Characteristics of Studies Reviewed

The methodological features of the 14 reviewed studies are reported in Table 1. As

shown, 8 were controlled studies, of which 4 featured random assignment of either classes or schools to intervention and control conditions. A further 6 were uncontrolled studies. Self-reported bullying and victimization (see Table 2) were included as outcome measures in all but one study and are consequently the focus of our analysis. Table 3 summarizes the intervention components of the whole-school programs examined in our research synthesis. Whole-school programs by their nature have multiple components that usually operate simultaneously at different levels in the school context. As shown in Table 3, the intervention programs comprised components at a minimum of three program levels to a maximum of five levels.

Overall Program Effects

The data from the original studies were generally of three types: (a) percentages of participants reporting involvement in bullying before and after the intervention (10 studies), (b) scores on a bullying or victimization scale (3 studies), and (c) narrative descriptions (2 studies; one study contained both narrative and quantitative data). Table 2 lists the results of the 14 studies as r values.[1] To compute these values, standardized mean differences (i.e., z values) were calculated for percentage and score data from the original studies and then using *MetaWin 2.0* were transformed into r values with the formula, $r = Z/\sqrt{N}$ (Rosenthal, 1994). Ns within studies having unequal samples from pretest to posttest were averaged, and finally, all aggregated r values were weighted by sample size.

In 3 of the 14 studies, antibullying programs were implemented in both primary and secondary schools. Although the available data were not sufficient for a year-by-year calculation of age or grade effect, separate effects are reported for primary and secondary schools where appropriate. (It should be remembered, however, that the transition to secondary school does not come at exactly the same age in each country. Therefore, this distinction is, at best, a rough one.) For purposes of comparing the overall effectiveness of antibullying programs, data from programs in primary and secondary

schools within the same study are treated as constituting different programs.

Several studies reported effects for intervention conditions in addition to the basic pretest/posttest evaluation. Three such studies included a follow-up evaluation at intervals ranging from 5 months to 1 year post-program. One presented effects of an intervention at two levels of program implementation. One study included a total of three pretest/posttest evaluations, each with a different cohort of students over 3 consecutive years. One reported bullying and victimization outcomes for different time frames (i.e., within the last 5 days and within the last 2 months). For these studies, an average effect was calculated for all intervention conditions within the given study (weighted by Ns within conditions) and was used as one of the primary statistics in the analyses. In some studies, this could represent a deceptive average because it might combine conditions considered optimal and less than optimal. For example, in the study by Stevens, De Bourdeaudhuij, and Van Oost (2000), there were two implementation conditions—with and without intensive support and consultation by professionals. In some studies, it is argued that the effects of the intervention become apparent only after they had penetrated school culture and therefore might be apparent at follow-up measurement but not immediately after treatment. To test the limits of the program effectiveness represented in the data, Table 2 presents the r values corresponding to the single best intervention effects.

Table 4 summarizes the effects of the intervention programs on self-reported victimization and bullying across five levels of effect size, using categories proposed by Cohen (1988). Effects fell almost exclusively in the categories of small, negligible, and negative. Only 7% of the total effects reported in the table (i.e., only one condition in one study) were categorized as medium and none as large. Looking at the effects corresponding to the averaged intervention conditions for the victimization outcomes, 93% were negligible or negative. For self-reported bullying outcomes, 92% were negligible or negative. Considering only the best intervention effects, 67% of stud-

Table 1
Study Characteristics

Author (year)	Location	Total n	Grades (ages) of Participants	n schools	Months to Posttest (follow-up)	Properties of Measures Reported
Controlled Studies With Random Assignment to Groups						
Ciucci & Smorti (1998)	Italy	287	1-3	6	10	reliability
Melton et al. (1998)	South Carolina, USA	6264	4-8	18	12 (24)	reliability, validity
Stevens et al. (2000)	Belgium	primary: 290 secondary: 526	(10-16)	18	19	reliability
Rosenbluth & Sanchez (2002)	Texas, USA	1406	5	12	3 (8)	reliability, validity
Controlled Studies With Nonrandom Assignment to Groups						
Whitney et al. (1994)	England	primary: 2311 secondary: 5998	(8-16)	27	24	none
Twemlow et al. (1999)	Kansas, USA	542	1-5	2	24	reliability, validity
Alsaker & Valkanover (2001)	Switzerland	319	K	16	4	none
Rahey & Craig (2002)	Canada	530	K-8	2	4	reliability
Uncontrolled Studies						
Munthe (1989)	Norway	3800	(10-16)	37	36	none
Olweus (1993)	Norway	2500	(13-14)	42	20	reliability
Pepler et al. (1994)	Canada	1047	(8-14)	4	18	none
Hanewinkel & Knaack (1997)	Germany	primary: 3246 secondary: 7574	3-12	37	24	none
Peterson & Rigby (1999)	Australia	708	7, 9-11	1	24	reliability
Salmivalli et al. (2003)	Finland	761	(9-11)	16	12	reliability

Table 2
Summary of Results

	Intervention Group Effects		Control Group Effects		Program Effects		Integrity Data
	Bullying M^a (best)[b]	Victimization M (best)	Bullying M	Victimization M	Bullying M	Victimization M	
Ciucci & Smorti (1998)	.06 (.10)	.08 (.24)	.02	.01	.06	.09	none
Melton et al. (1998)	.03 (.05)	.09 (.10)	.04	.06	-.02	.02	interviews, activity logs
Stevens et al. (2000) primary:	.02 (.07)	.09 (.20)	-.13	.05	.08	.12	none
secondary:	-.04 (.04)	-.01 (.04)	-.02	.00	.05	-.01	
Rosenbluth & Sanchez (2002)	—	—	—	—	-.05	-.03	none
Whitney et al. (1994) primary:	.05	.09	-.32	-.17	—	—	interviews, activity logs
secondary:	.00	-.05	.07	.02	—	—	
Twemlow et al. (1999)	—	.08	—	.11	—	.00	none
Alsaker & Valkanover (2001)[c]	.03	.01	.12	.04	—	—	interviews, questionnaires
Rahey & Craig (2002)	—	—	—	—	.00	.00	activity logs
Munthe (1989)	.03	.02					none
Olweus (1993)	.29	.33					interviews, activity logs
Pepler et al. (1994)	-.07 (-.05)	.01 (.08)					questionnaires, interviews
Hanewinkel & Knaack (1997) primary:	.02	.05					none
secondary:	.07	.02					
Peterson & Rigby (1999)		.01					none
Salmivalli et al. (2003)	09 (.15)	.09 (.13)					questionnaires

Note. Missing data in the table are indicated by a dash (—). These data were not reported in the original studies.
[a]M = the average effect for all intervention conditions within the study. [b]Best = the single best intervention effect within the study. [c]Data for this study are peer nominations, not self-reports.

Table 3
Program Components

Study	School						Parent			Classroom			Peers	Individuals
	Anti-bullying Policy	Increased Supervision	Playground reorganized	Information	Anti-bullying Committee	Staff Training	Information	Involved in anti-bullying activities	Targeted interventions	Rules	Curricular activities	Social skills training	Peer-led interventions	Targeted interventions for bullies and victims
Munthe (1989)	•						•			•			•	•
Olweus (1993)		•	•	•		•	•	•	•	•	•		•	•
Whitney et al. (1994)	•	•	•	•			•	•	•	•	•		•	•
Pepler et al. (1994)	•	•	•	•							•		•	•
Hanewinkel & Knaack (1997)		•	•	•		•	•		•	•	•			
Ciucci & Smorti (1998)	•			•		•	•							
Rahey & Craig (2002)		•		•		•	•		•	•	•		•	•
Melton et al. (1998)		•	•	•	•				•		•	•		
Twemlow et al. (1999)	•			•							•	•	•	•
Peterson & Rigby (1999)	•			•	•	•	•					•	•	•
Stevens et al. (2000)	•			•		•	•			•	•	•	•	•
Alsaker & Valkanover (2001)				•		•	•			•			•	•
Salmivalli et al. (2003)	•			•						•	•	•		•
Rosenbluth & Sanchez (2002)	•			•		•	•				•	•		•

ies revealed small effect sizes for victimization outcomes and the remaining 33% revealed negligible effects. For bullying outcomes, 33% of reported effects were small, and 67% negligible or negative. The study designs permitted comparison of intervention and control groups (i.e., "program conditions") for seven antibullying programs. This is arguably the most rigorous test of program effectiveness permitted by our analysis. Under these more discerning evaluation conditions, 86% of victimization outcomes were negligible or negative and the remaining 14% of reported effects were positive (albeit small). For self-reported bullying, 100% of the reported effects were negligible or negative.

Table 5 provides additional detail on the self-reports. It is important to note that the percentages reported in this table should not be compared across studies, because they are not based on the same question and do not necessarily pertain to the same time interval (e.g., previous week, previous month, previous school term). Therefore, each posttest percentage is properly compared only with the corresponding pretest percentage from the same study.

Treatment integrity. As shown in Table 2, 7 of the 14 studies incorporated some systematic verification that the intervention was delivered with integrity. Two levels of program monitoring (categorized as present or absent) were cross-tabulated with intervention effects (averaged) for victimization and bullying outcomes (categorized as statistically significant improvement vs. no improvement). This tabulation revealed that programs with a monitoring component yielded more positive outcomes than programs lacking any formal monitoring procedures. The link between program monitoring and outcomes reached statistical significance for victimization self-reports, $\chi^2 = 5.33$, $p = .02$), but it fell short of significance for bullying self-reports, $\chi^2 = 2.86$, $p = .09$.

Age effects. School level (primary and middle schools vs. secondary schools) was cross-tabulated with victimization and bullying outcomes (statistically significant improvement vs. no improvement). Although only programs in primary and middle schools (Grades 1 through 8 or approximately ages 7 to 14

years) yielded any positive outcomes and programs in secondary schools ($N = 4$) yielded only null effects, the trends did not reach statistical significance for either bullying, $\chi^2 = .94$, $p = .33$, or victimization, $\chi^2 = 1.61$, $p = .21$.

Discussion

Before contemplating the results and their implications, it is important to comment on the adequacy of the available evidence. Although most of the studies contained the core elements of the whole-school approach introduced by Olweus (1993), implementation of the approach varied considerably. This makes the results difficult to synthesize. Moreover, although there are many commonalities in the outcome measures, there is not enough similarity among them to compare the results with the confidence that would emerge in comparing data from identical measures.

A number of drawbacks in the studies themselves further complicate interpretation of this literature. Control conditions were absent in many studies, and, in a number of others, schools self-selected for experimental and control conditions, with the more motivated schools (or those motivated to begin the program earlier) opting for the experimental condition. Only some of the studies incorporated systematic procedures to ensure that the planned interventions were implemented with integrity, and in several cases implementation of some of the program components was optional for the schools involved.

The most common outcome measures were self-reports of victimization and bullying. These are not strictly comparable across studies because they refer to incidents occurring at different periods of time (i.e., the past week, the past month, the past 3 months). In any case, recent research has illustrated that self-reports of bullying and victimization, although not necessarily inaccurate, do not correspond to information about bullying and victimization obtained from peers or teachers or from observations (Pellegrini & Bartini, 2000). An added confounding issue is the effect that sensitization to information about bullying can have on students' reports of their experiences of bullying. Antibullying programs obviously

Table 4
Summary of Effect Sizes

Study condition	Outcome	Large ≥ .50	Medium .30–.49	Small .10–.29	Negligible .00–.09	Negative < .00
Intervention conditions: Average	Victimization	0	7	0	80	13
	Bullying	0	0	8	77	15
Intervention conditions: Best	Victimization	0	0	67	33	0
	Bullying	0	0	33	50	17
Program conditions in controlled studies	Victimization	0	0	14	57	29
	Bullying	0	0	0	67	33

Note. Figures indicate percentage of studies achieving respective levels.

Table 5
Changes in Study Outcomes

Study	Intervention		Control	
	Bullying	Victimization	Bullying	Victimization
Ciucci & Smorti (1998)	51.3 → 47.3	47.3 → 41.7	52.7 → 51.3	46.7 → 47.7
Melton et al. (1998)	24 → 20 → 25	25 → 19 → 17	19 → 22	24 → 19
Rosenbluth & Sanchez (2002)	10.6 → 17.0 → 20.8	40.8 → 36.7 → 30.1	11.2 → 17.8 → 13.7	47.5 → 34.7 → 28.7
Whitney et al. (1994) primary:	37.7 → 34.5	59.3 → 53.2	34.5 → 56.9	52.2 → 63.8
secondary:	26.0 → 25.2	26.8 → 29.8	34.6 → 29.9	42.9 → 41.5
Alsaker & Valkanover (2001)	12.5 → 13	12 → 10.5	9 → 10.5	7 → 12
Munthe (1989)	2.4 → 3.0	3.8 → 4.4		
Olweus (1993)	37.3 → 19	34 → 14.3		
Pepler et al. (1994) past 2 months:	7 → 9	12 → 15		
past 5 days:	16 → 21	28 → 23		
Salmivalli et al. (2003)				
low program integrity:	12.9 → 10.0	13.5 → 10.5		
high program integrity:	12.9 → 5.1	13.5 → 6.5		

Note. Figures indicate the percentage of children reporting involvement in bullying. The first value in each cell is the pretest percent; the second is the posttest percent. When follow-up data were available, they appear as the third value in the cell.

increase awareness of the phenomenon, which may cause students to more frequently report bullying incidents at school and essentially mask a positive effect of the whole-school program.

Despite these limitations some conclusions can be drawn with caution. It is clear that the whole school approach has led to important reductions in bullying in a number of cases, but the results are simply too inconsistent to justify adoption of these procedures to the exclusion of others. The widespread enthusiasm for the whole-school approach, and its enactment into law in some jurisdictions, can be based only on the perceived urgent need to intervene and on the few studies indicating success. The dramatic success of the Olweus program in Norway has not been replicated elsewhere. There are a few instances of significant improvement following program implementation, though not nearly as striking, and there are many nonsignificant findings and some results opposite to the expected direction. This pattern of modest results is consistent with other reviews of school-based primary prevention programs (e.g., Howard, Flora, & Griffin, 1999). The question of the clinical significance of the findings cannot be answered readily from the data at hand because such information is not available in the original studies. This approach may be useful in future research, but would require a considerable investment to establish, for example, thresholds at which a reduction in bullying and victimization can be deemed to have reduced the risk of long-term maladjustment or stable patterns of externalizing or internalizing problems.

We propose several possible explanations for the lack of consistency in the evaluation findings. One possibility is that Olweus's impressive success relates to the high quality of Scandinavian schools, which have small classes and well-trained teachers, together with the well-ingrained Scandinavian tradition of state intervention in matters of social welfare. The smaller but nonetheless successful results in Finland and Italy could also relate to the quality of the schools there. If this is the case, more attention might be devoted to interventions that address the broader aspects of school climate, such as the general atmosphere and

interpersonal relationships of the school, as discussed in some recent writings on bullying by Fernández (2001). The success of the Olweus program may be related in part to its historical context, perhaps making it a unique and unreplicable case. The program was introduced into schools on a national scale in the wake of several highly publicized suicides that were linked publicly to bullying (Olweus, 1993). It seems plausible that this could have increased the seriousness and urgency with which school officials and students invested themselves in the initiative, which subsequently had a positive effect on the program outcomes.

Another interpretation of general findings in this article is that the results, inconsistent as they are, reflect a reasonable rate of return on the investment inherent in low-cost, nonstigmatizing primary prevention programs. This might be of particular importance because bullying and other forms of aggression are probably highly refractory to intervention of any type, and the importance of small, significant findings should not be dismissed (see Schneider, 1992, 1993). The possibility that Olweus and his Norwegian colleagues developed a unique package of intervention components that is ineffective when diluted or modified cannot be ruled out. The benefits of implementing interventions with fidelity versus adapting them to local conditions have been debated (e.g., Dane & Schneider, 1998). Most of the studies reviewed herein have entailed substantial modifications to the original program. Although such changes may be justified, they typically are not described in the research reports in sufficient detail for readers to ascertain what possible influence these adaptations might have had on outcomes. Additionally, when researchers inadequately describe the interventions and implementation procedures in their reports, the transportability of programs is undermined. This leaves school professionals unable to replicate the interventions with fidelity.

A tabulation of best effects across studies suggests that the whole-school approach may have more potential than the evaluation data indicate. In some studies, effectiveness was not apparent until follow-up; in others, effectiveness was attenuated at follow-up (see

Table 5). It is impossible to identify the ingredients of a successful implementation in these studies. Consistent, methodical effort is needed to determine exactly which components or conditions are key to making the approach effective.

In conclusion, only a cautious recommendation can be made that whole-school antibullying interventions be continued until they are evaluated further. This recommendation is based not on solid evidence that the programs work, but rather on the logical links between programs and theories about the origins of bullying and because in some instances (and under the most favorable conditions researchers have been able to contrive) they have been effective. This is not to say that any other form of intervention appears to be more effective than these. Indeed, there is no evidence that other forms of intervention are superior to the whole-school approach in dealing with bully-victim problems.

Implications for Research

Future research should include rigorous monitoring of program implementation in the schools to ensure that schools are actually receiving the intended programs. The synthesis in this article has revealed that program monitoring is linked to at least one important program outcome (i.e., victimization). Researchers should collect data on outcomes from other sources in addition to students' self-reports, such as observations from teachers, classmates, administrators, and even parents. This is particularly important in the light of findings on the divergence of perceptions among those who witness bullying (Pellegrini & Bartini, 2000).

Implications for Practice

Despite the limited empirical support for the effectiveness of antibullying programs, there is not sufficient evidence to conclude that these programs should be abandoned. On the other hand, there is no basis for psychologists and educators to participate in a campaign advocating the resolution of bully/victim problems by means of the whole-school approach to the exclusion of any other modality. The overarching message is that intervention can

succeed, but not enough is known to indicate exactly how and when. Psychologists should be proactive in promoting carefully evaluated interventions in which the whole-school approach is implemented with precision and compared with other potentially useful interventions.

Footnote

[1]The two studies reporting effects only with narrative descriptors yielded nonsignificant results, and these effects are estimated as $r = .00$ in the table.

References

References marked with an asterisk indicate studies included in the research synthesis.

*Alsaker, F. D., & Valkanover, S. (2001). Early diagnosis and prevention of victimization in kindergarten. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 175-195). New York: Guilford Press.

*Ciucci, E., & Smorti, A. (1998). Il fenomeno delle pretonenze nella scuola: problemi e prospettive di intervento. [The phenomenon of bullying in school: Problems and prospects for intervention]. *Psichiatria dell'infanzia e dell'adolescenza, 65*, 147-157.

Cohen, J. (1988). *Statistical power for behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Craig, W. M., Pepler, D., & Atlas, R. (2000). Observations of bullying in the playground and in the classroom. *School Psychology International, 21*, 22-36.

Dane, A., & Schneider, B. H. (1998). Integrity in primary prevention programs: Are implementation effects out of control? *Clinical Psychology Review, 18*, 23-45.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755-764.

Farrington, D. (1993) Understanding and preventing bullying. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 17, pp. 381-458). Chicago: University of Chicago Press.

Farrington, D. P., & Hawkins, J. D. (1991). Predicting participation, early onset, and later persistence in officially recorded offending. *Criminal Behavior and Mental Health, 1*, 1-33.

Fernández, I. (2001). *Prevención de la violencia y resolucón de conflictos: El clima escolar como factor de cualidad.* [*Violence prevention and conflict resolution: School climate as a quality factor*]. Madrid: Narcea.

Glover, D., Gough, G., Johnson, M., & Cartwright, N. (2000). Bullying in 25 secondary schools: Incidence, impact and intervention. *Educational Research, 42*, 141-156.

*Hanewinkel, R., & Knaack, R. (1997). *Mobbing: Gewaltpraevention in schulen in Schleswig-Holstein.* [*A preventive anti-bullying program in the schools of Schleswig-Holstein*]. Kiel, Germany:

Gemeindeunfallversischerungsverband Schleswig-Holstein.

Howard, K. A., Flora, J., & Griffin, M. (1999). Violence-preventions programs in schools: State of the science and implications for future research. *Applied and Preventative Psychology, 8*, 197-215.

Kumpulainen, K., Raesaenen, E., & Henttonen, I. (1999). Children involved in bullying: Psychological disturbance and the persistence of the involvement. *Child Abuse & Neglect, 23*, 1253-1262.

*Melton, G. B., Limber, S. P., Cunningham, P., Osgood, D. W., Chambers, J., Flerx, V., Henggeler, S., & Nation, M. (1998). *Violence among rural youth: Final report to the Office of Juvenile Justice and Delinquency Prevention*. Clemson, SC: Clemson University, Institute on Family and Neighborhood Life.

*Munthe, E. (1989). Bullying in Scandinavia. In E. Roland & E. Munthe (Eds.), *Bullying: An international perspective* (pp. 66-78). London: David Fulton.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association, 285*, 2094-2100.

*Olweus, D. (1993). *Bullying at school: What we know and what we can do.* Oxford, UK: Blackwell.

Olweus, D. (1994). Bullying at school: Long-term outcomes for the victims and an effective school-based intervention program. In L. R. Huesmann (Ed.), *Aggressive behavior: Current perspectives* (pp. 97-130). New York: Plenum.

Olweus, D. (1997). Bully/victim problems in school: Facts and intervention. *European Journal of Psychology of Education, 12,* 495-510.

Ortega, R., & Lera, M. (2000). The Seville Anti-Bullying in School Project. *Aggressive Behavior, 26*, 113-123.

Pellegrini, A. (2001). The roles of dominance and bullying in the development of early heterosexual relationships. *Journal of Emotional Abuse, 2,* 63-73.

Pellegrini, A., & Bartini, M. (2000). An empirical comparison of methods of sampling aggression and victimization in school settings. *Journal of Educational Psychology, 92*, 360-366.

*Pepler, D. J., Craig, W. M., Ziegler, S., & Charach, A. (1994). An evaluation of an anti-bullying intervention in Toronto schools. *Canadian Journal of Community Mental Health, 13*, 95-110.

*Peterson, L., & Rigby, K. (1999). Countering bullying at an Australian secondary school with students as helpers. *Journal of Adolescence, 22*, 481-492.

*Rahey, L., & Craig, W. M. (2002). Evaluation of an ecological program to reduce bullying in schools. *Canadian Journal of Counselling, 36*, 281-296.

Roberts, W. B., & Morotti, A. A. (2000). The bully as victim: Understanding bully behaviors to increase the effectiveness of interventions in the bully-victim dyad. *Professional School Counseling, 4*, 148-155.

*Rosenbluth, B., & Sanchez, E. (2002). *The Expect Respect project: Preventing bullying and sexual harassment in the U. S. elementary schools.* Manuscript submitted for publication.

Rosenthal, R. (1994). Parametric measures of effect sizes. In H. Cooper & L. V. Hedges (Eds.), *Handbook of research synthesis* (pp. 431-244). New York: Sage Foundation.

Salmivalli, C. (1999). Participant role approach to school bullying: Implications for intervention. *Journal of Adolescence, 22*, 453-459.

*Salmivalli, C., Kaukiainen, A., & Voeten, M. (2003). *Antibullying intervention: Implementation and outcome.* Manuscript submitted for publication.

Schneider, B. H. (1992). Didactic methods for enhancing children's peer relations: A quantitative review. *Clinical Psychology Review, 12*, 363-382.

Schneider, B. H. (1993). Social skills training and the aggressive child: On the value of small paradoxical gain. *Exceptionality Education Canada, 3*, 195-203.

Smith, P. K., & Brain, P. (2000). Bullying in schools: Lessons from two decades of research. *Aggressive Behavior, 26*, 1-9.

Smith, P. K., & Myron-Wilson, R. (1998). Parenting and school bullying. *Clinical Child Psychology and Psychiatry, 3*, 405-417.

Sourander, A., Helstelae, L., Helenius, H., & Piha, J. (2000). Persistence of bullying from childhood to adolescence: A longitudinal 8-year follow-up study. *Child Abuse & Neglect, 24*, 873-881.

*Stevens, V., De Bourdeaudhuij, I., & Van Oost, P. (2000). Bullying in Flemish schools: An evaluation of antibullying intervention in primary and secondary schools. *British Journal of Educational Psychology, 70,* 195-210.

Sutton, J., & Smith, P. K. (1999). Bullying as a group process: An adaptation of the participant role approach. *Aggressive Behavior, 25*, 97-111.

*Twemlow, S. W., Fonagy, P., Sacco, F. C., Gies, M. L., Evans, R., & Ewbank, R. (1999). *Creating a peaceful school learning environment: A controlled study of an elementary school intervention to reduce violence.* Unpublished manuscript, Menninger Clinic, Child and Family Center, Topeka, KS.

*Whitney, I., Rivers, I., Smith, P. K., & Sharp, S. (1994). The Sheffield Project: Methodology and findings. In P. K. Smith & S. Sharp (Eds.), *School bullying insights and perspectives* (pp. 20-56). London: Routledge.

J. David Smith is an Associate Professor of educational counseling at the Faculty of Education, University of Ottawa, Canada. He received his PhD in counseling psychology in 1998 from McGill University. His research activities are focused on evaluating and improving the effectiveness of school-based antibullying programs.

Barry H. Schneider is Associate Dean (research) and Professor of Psychology at the Faculty of Social Sciences, University of Ottawa, Canada, where he has taught since 1981. He received his PhD in counseling psychology from the University of Toronto in 1977. His major research interest is the interpersonal relationships of children with both externalizing and internalizing problems, including interventions designed to improve their relationships.

Peter K. Smith is Professor of Psychology and Head of the Unit for School and Family Studies at Goldsmiths College, University of London; he is a Fellow of the British Psychological Society. He is co-author of *Understanding Children's Development* (Blackwell, 1988, 1991, 1998, 2002), and co-editor of *School Bullying: Insights and Perspectives* (Routledge, 1994), *Tackling Bullying in Your School: A Practical Handbook for Teachers* (Routledge, 1994), *The Nature of School Bullying* (Routledge, 1999), The Family System Test (Routledge, 2001), and the *Blackwell Handbook of Childhood Social Development* (Blackwell, 2002), editor of *Violence in Schools: The Response in Europe* (Routledge, 2002), and co-editor of *Bullying in Schools: How Successful Can Interventions Be?* (Cambridge University Press, 2004).

Katerina Ananiadou has a background in psychology. She is a research officer at the Institute of Education, University of London, currently working in the field of adult basic skills within the National Research and Development Centre for adult literacy and numeracy.