

Received: 2004.10.04
Accepted: 2005.02.17
Published: 2005.07.01

Creating a peaceful school learning environment: the impact of an antibullying program on educational attainment in elementary schools

Authors' Contribution:

- A** Study Design
- B** Data Collection
- C** Statistical Analysis
- D** Data Interpretation
- E** Manuscript Preparation
- F** Literature Search
- G** Funds Collection

Peter Fonagy^{1ACDEF}, Stuart W. Twemlow^{2ABDEFG}, Eric Vernberg^{3ABDEFG},
Frank C. Sacco^{4BCE}, Todd D. Little^{3BDEF}

¹ University College London, England

² Menninger Clinic, Houston, TX, U.S.A.

³ University of Kansas, Kansas, KS, U.S.A.

⁴ Community Services Institute, Springfield, MS, U.S.A.

Source of support: Financial support from the Menninger Foundation Endowment for the Child and Family Center

Summary

Background:	The impact of a bullying and violence prevention program on education attainment was studied in five elementary schools (K-5), over a 5-year period.
Material/Methods:	A multiple baseline design was used and academic attainment test scores of 1,100 students were monitored before and after the introduction of the program across the school district. This sample was contrasted with an equivalent control sample of 1,100 students from the school district who attended schools that did not join the program.
Results:	Program participation was associated with pronounced improvements in the students' achievement test scores. Notable reductions in the scores of those students who left schools with active programs were also observed.
Conclusions:	This simple, low-cost anti-violence intervention, involves all those who work in schools, not just students. It appears to significantly benefit educational performance of children in the participating elementary schools. The program focuses attention on the interaction between the bully, victim and audience of bystanders who are seen as pivotal in either promoting or ameliorating violence. Buy in to the philosophy by teachers & administration is high, because the format allows each school to create materials with its own personal stamp, and since there is no classroom curriculum add on, the burden to teachers is vastly reduced. Psychiatrists who work with schools could easily assist a school to put the program in place as part of their consultation work.
key words:	bullying • educational attainment • schools and violence prevention

Full-text PDF: http://www.MedSciMonit.com/pub/vol_11/no_7/6440.pdf

Word count: 5333

Tables: 3

Figures: -

References: 48

Author's address: Stuart W. Twemlow MD, P.O. Box 809045, 2801 Gessner Drive, Houston, TX 77280-9045, U.S.A.
e-mail: stwemlow@aol.com

BACKGROUND

Violence prevention has become an increasingly important priority in the U.S. [1] and abroad. Particularly in school settings, violence among peers is quite problematic [2-4] and many schools appear to have a propensity for generating informal social norms that predispose them to violence [5-7].

School-based programs generally attempt to challenge and change these norms and reduce the level of various forms of aggression among peers. Success in accomplishing these aims varies. A meta-analysis of school-based violence-prevention programs [8], found significant reductions in aggressiveness for researcher-initiated programs, but much smaller effects for school-initiated programs, with high-risk youth showing greater reductions in aggressive behavior across all programs than non-high-risk youth. Another meta-analysis [9] of 44 randomized-controlled trials of programs focused on children identified as at-risk for aggressive behavior. Although reductions in aggressiveness were consistently shown, the effect sizes were small. The reductions were more striking in programs addressing mixed-sex groups rather than boys alone. The degree of violence among peers has also been demonstrated to be a key mediator of the relations between family factors and engagement in acts of serious delinquency or violence [10]. Overall, these studies suggest that intervention-related changes in peer aggression are possible and may have benefits for the overall academic mission of schools.

Elementary school may be an especially desirable setting for prevention efforts. Developmental studies repeatedly emphasize that the early onset of aggression and violence predicts a child's future risk for violence [11]. Children's beliefs about aggression and maladaptive attribution biases concerning the hostile intent of others tend to crystallize between the ages of 6 and 12 [12]. Middle childhood is an important time for integrating the emotional, cognitive, and behavioral skills needed to develop social competence [13]. Moreover, elementary schools can provide relatively unencumbered access to high-risk populations and relevant outcome data such as educational attainment [14,15].

In our view, educational attainment merits consideration as an important outcome of violence prevention programs for a number of reasons. Violence towards elementary-aged children potentially interferes with the learning process in many ways. Children who are repeatedly targeted by peers as victims of aggression show elevated symptoms of depression and anxiety, which can interfere with concentration and task completion [16-18]. Exposure to interpersonal violence in community settings is related to poor academic performance and this relation appears mediated by depressive symptoms and disruptive behaviors that often appear in the aftermath of aggression [19]. Anger, embarrassment, or fear generated by aggressive acts among peers, directs children's energies toward revenge or avoidance rather than academic work. Ongoing conflicts among children also disrupt instructional time. Among children who perpetrate violence toward peers, educational problems are common. Highly aggressive youth, for example, are more likely than less aggressive peers to have been retained in one or more grades [20].

The combination of poor academic achievement and early conduct problems predicts criminal involvement across the lifespan [21-24]. Behavioral problems have been linked to reading and language difficulties [25-27], with the association of language problems and behavioral disorders apparently mediated by reading difficulties [28]. Conversely, educational attainment appears to function as a protective factor - an indicator of so-called resilient children [29-31].

Enhanced educational attainment is a concern, not just for individuals involved in violence as aggressors or as victims, but also for the entire community. School violence that adversely affects the educational process depletes the value of the human capital of a town, state, and nation. The associations between the educational attainment of a particular school and the school's capacity to generate an atmosphere that is likely to promote healthy socialization of its students are well known [32].

Educational attainment is a desirable outcome measure because of its strong psychometric properties [33], including sensitivity to change and resistance-response bias. In this regard, educational attainment offers a sound index of enduring conditions in the school. Focusing on educational attainment in schools involved in violence prevention may also overcome resistance to preventive interventions on the part of teachers, principals, administrators, and board members by creating a common cause between mental health professionals and educators [34,35].

To our knowledge, large-scale studies examining the impact of elementary-school violence-prevention programs on children's educational attainment have not been conducted. A number of studies have demonstrated promising effects for violence-prevention programs in elementary schools, primarily for changes in violence-related cognitions and teacher- or child-reports of aggressive behavior. Programs producing these effects share features such as instruction and practice in identifying, managing, and coping with negative emotions, appeals to moral reasoning, teaching non-violent alternatives to conflicts, enhancing negotiation, thinking, and decision-making skills (e.g., envisioning the consequences of aggressive acts). The current study utilizes a school-based violence-prevention program with these features, but focuses on educational attainment as the primary outcome.

The Creating a Peaceful School Learning Environment (CAPSLE) initiative started in Topeka, Kansas, with an intervention in a single elementary school in 1994 [36]. CAPSLE is an open social-systems, psychodynamically informed, intervention that integrates several approaches addressing the dialectical relationship between the victimizer (the bully), the victim, and the bystanding "audience" whose complex interaction with the bully and victim help facilitate or inhibit the victimizing process [37,38]. When violence becomes a serious and pervasive problem in the school, the entire school environment is considered, from this viewpoint, to be dysfunctional. This pathological social system is assumed to occur not only between students but also between members of the school community. Thus, the CAPSLE program addresses children as well as school employees, volunteers, and parents. Schools with high levels of bully-victim problems often lack a plan to address the overall school social climate, and rely heavily on punitive or coercive discipline strategies [37,38].

The CAPSLE program utilizes four primary components. First, a positive climate campaign uses reflective classroom discussion, counselor-led lessons, posters, magnets, book-marks, and other devices to encourage a shift in language (and thinking) of all students and personnel. These language tools help identify and resolve problems that occur when coercive power-dynamics dominate the school environment [38]. For example, children help each other resolve issues without adult participation. Such effects are observed as they share playground equipment peacefully, and do not push and jostle in the lunch line. Second, a classroom management plan assists teachers' discipline skills by focusing on understanding and correcting the root problems rather than on punishing and criticizing the behavior. For example, a behavior problem in a single child is conceptualized as a problem for the whole class who, often unwittingly, participate in bully victim or bystander roles. Scapegoating is thereby reduced and insight into the meaning of the behavior becomes paramount. Third, a physical education program derived from a combination of role-playing, relaxation, and self-defensive techniques, teaches children skills to handle victimization and bystander behavior. This program helps children protect themselves and others with non-aggressive physical and cognitive strategies. For example, enacting bully-victim-bystander roles provides students with alternative actions to fighting. Learning ways to physically defend oneself (e.g., when grabbed, pushed, or punched) coupled with classroom discussion, teaches personal self control as well as respect and helpfulness toward others. Fourth, schools may put in place one or two support programs: peer mentorship or adult mentorship. These relationships provide additional containment and modeling to assist children in mastering the skills and language to deal with power struggles. For example, mentors instruct children in refereeing games, resolving playground disputes, and the importance of helping others.

From 1996 through 2000, the program extended out to five schools in the school district. The gradual rollout of the program enabled us to contrast the performance of children attending CAPSLE schools with that of children from other schools in the same district. Tracking the movement of children between schools also permitted a comparison of children's educational attainment while in a CAPSLE school with their performance (a) prior to the school joining the program, (b) their performance prior to their individually joining a participating school or (c) when leaving the school to move to another school where the program was not available. This study aims to identify whether improvements in educational attainment are associated with attendance at schools where the program was implemented. In terms of hypotheses, we expected to find pervasive support for the benefits of the CAPSLE program in terms of enhanced educational attainment outcomes.

MATERIAL AND METHODS

The CAPSLE program was rolled out in a step-wise fashion in one school district over the period 1994-1999. The program was started in a single school, with an additional school incorporated in 1996 and three further schools in 1999. The initial school was identified as a particular problem school and the second was a demographically matched school. The three schools included in 1999 were selected

randomly from 9 schools that volunteered for the program [39]. The stepwise rollout presents multiple baselines [40] because, other than the first school, program implementation was not contingent on school climate. Thus, changes associated with the timing of the program's introduction would not be confounded with changes caused by the program itself.

School staff implemented the program with support from a consultation team led by the first three authors, teachers, counselors, and building administrators took primary responsibility for the positive climate campaign and the classroom management plan. The physical education program included 12 sessions delivered once weekly during regular physical education time. These sessions were co-taught by the physical education teacher and a martial arts specialist. School counselors or social workers coordinated and supervised the peer mentor or adult mentor program. All interventions were described in a detailed manual designed to enhance program fidelity and replicability [41].

Participants

A total of 2,206 students participated in this study, all of which provided written parental consent. According to the data supplied by the school district we found that on the average 98.2% (97.2-98.9%) of those registered for the test had valid records. There was a tendency toward slightly fewer missing cases in more recent years despite increased numbers taking the test. For the CAPSLE group, standardized achievement test scores were available for 1,106 students who attended a school in which the program was implemented for one or more years from 1996 through 2000. Achievement scores were available in 1996 for 156 (14.1%) students, 1997 for 198 (17.9%) students, 1998 for 236 (21.3%) students, 1999 for 212 (19.1%) students, and 2000 for 304 (27.5%) students. There were 587 (53.1%) male and 519 (46.9%) female students. The majority of students were white (58.4%), 23.3% were African American, 14.2% were Hispanic, 3.1% Native Americans and 1.1% Asians. In 1996, the majority of the students received free or reduced-price lunches (59.4%), and ages ranged from 7 to 14 years ($n=9.7$).

A matched comparison group was constructed by randomly selecting from non-CAPSLE students whose gender, ethnicity, age, lunch-program status, and test year for which test scores were available matched the program student's characteristics. Using this procedure, 1,100 control students were recruited (complete demographic information was not available for six program participants, so no matched controls could be recruited for these six).

In addition to the comparison with the control group, we compared participants' achievement test scores before and after program implementation. To increase the number of years on which such comparisons could be based, any available test scores for the years of 1993, 1994, and 1995 were obtained for students who later participated in the program. Any available test scores for the same years were obtained for the matched control students. Usually because of family relocation, a number of students moved out of CAPSLE schools during the study period into other elementary schools in the same school district. The number of

students moving out of CAPSLE schools after one year was 118 (11%), after two years was 6 (less than 1%), and after three years was 76 (7%). The average number of students per year from 1997 through 2000 who moved from a program school in the previous year to a non-program school in the current year was 50.0 (18%) with a range from 0 (in 1997) to 105 (in 1999).

Measures

Academic achievement was assessed using the Metropolitan Achievement Test (MAT7) [12], which is routinely administered to 3rd and 5th graders. The MAT is a comprehensive battery of tests designed to measure school achievement across the domains of reading (word recognition, reading vocabulary, reading comprehension), mathematics (concepts and problem solving and mathematical procedures), written language (composition, editing), science, social studies, research and thinking skills. The MAT7 was normed on a nationally representative sample of 180,000 children from kindergarten through 12th grade. Comparisons between test items and instructional goals suggested that the MAT7 demonstrated adequate content validity. It also had substantial construct validity evidenced by item difficulty analysis at different grade levels and clear discrimination among high- and low-scoring students. The reliability calculations for those scales utilized in the present study (i.e., basic battery, total reading, total math, and language were quite acceptable ($\alpha > 0.80$).

In this analysis, we looked at composite test scores (Basic Battery) as well as Reading, Language, and Math test scores. Depending on students' age, residency in the school district, and the achievement-testing schedule, test scores were available for one or more years from 1993 through 2000.

Analytic procedures

Students' achievement test scores were analyzed using the BMDP 5V analysis of variance program for unbalanced repeated measures models with structured covariance matrices [43]. This procedure uses the method of maximum likelihood to estimate parameters of a general class of models where the expected values of the responses are described as arbitrary linear functions of a set of regression parameters, and within-subject covariances are modeled as functions of a set of unknown covariance parameters. This allows time-varying covariates to be included into the multivariate analysis. Years for which test scores were available was treated as the within subject, repeated measures factor. The within-subject design was unbalanced because of the different numbers of years and combinations of years for which test scores were available for different students. A compound symmetry covariance structure, which specifies the same correlation between students' test scores for all pairs of years, was used to model individual differences in achievement. Model estimation was performed using maximum-likelihood estimation. Across all analyses, the estimated stability correlations between students' test scores across years were uniformly high (ranging from $r=0.69$ to 0.81), indicating that compound symmetry was a reasonable structure for modeling individual differences. This procedure allowed comparison of educational attainment based on multiple years of testing, rather than a single year.

RESULTS

Before turning to our tests of the effectiveness of the CAPSLE intervention, we first examined the influence of the demographic effects of age, gender, ethnicity (European American or Non-European American), and SES (i.e., free or reduced lunches). In these preliminary analyses, gender and ethnicity were treated as fixed covariates, while receipt of subsidized lunches and ages were treated as time-varying covariates. Separate analyses were performed for each of the four achievement tests. Estimates of the effects of each demographic variable were used to compute the marginal means. Specifically, marginal means were calculated using a linear combination of the model parameter estimates involved in the effect, holding all other parameters constant [44].

Table 1 shows the marginal means and tests of significance for the four achievement test scores. With a few exceptions, a quite consistent pattern of differences emerged across all four-achievement tests. Females were significantly higher than males on Basic Battery, Reading, and Language, but did not differ from males' scores on Math. Caucasian students were significantly higher than minority students on all four scales. Students who did not receive subsidized lunches were significantly higher than students who did on all measures. There were no significant effects for age on Basic Battery, Reading, and Language. On the age-normed Math test, younger students were higher than older students ($b=-0.81$, $X^2=12.20$, $df=1$, $p<0.001$). Because several of these demographic effects were significant, we controlled for them when testing the effects of participation in the CAPSLE program on student achievement.

Eight factors were analyzed to assess the effectiveness of participation in the CAPSLE program in facilitating academic achievement and to explore the possibility of differential program effectiveness for students in different demographic groups. The eight factors were (1) the year of measurement, (2) in a CAPSLE school (or not), (3) in CAPSLE school at year of test, (4) in CAPSLE school the year prior to test, (5) gender, (6) ethnicity, (7) subsidized-lunch status, and (8) age. Overall participation in the program was defined by whether a student attended a CAPSLE program school during a school year in which the program was implemented for any year from 1996 through 2000. To examine the extent to which CAPSLE attendance influenced performance for students who ever participated in the program, current year participation and previous year participation in the CAPSLE program were coded (1=yes, 0=no).

The model for testing the effectiveness of CAPSLE participation on achievement included main effects for gender, ethnicity, subsidized-lunch status, and age as well as year, overall participation (a fixed covariate), current year participation and previous year participation (time-varying effects nested within overall participation), the current year x previous year interaction effect (also nested within overall participation). The patterns of effects for the demographic variables were virtually identical to the patterns reported in Table 1.

Effects of program participation on achievement test scores

As before, separate analyses were performed for each of the four achievement tests. Estimated marginal means for each

Table 1. Test-score marginal means (and standard errors) for each demographic group.

Group	Test			
	Basic Battery	Reading	Language	Math
Gender				
Male; Nc = 1,169	46.50 (0.68)	45.89 (0.70)	46.19 (0.64)	48.48 (0.73)
Female; Nc = 1,034	48.30 (0.69)	47.93 (0.71)	50.10 (0.66)	47.46 (0.74)
χ^2 Male vs Female	4.80*	6.70**	25.72***	1.53
Ethnicity				
Caucasian; Nc = 1,287	52.51 (0.64)	51.58 (0.66)	51.85 (0.61)	52.96 (0.70)
Minority; Nc = 919	42.29 (0.73)	42.24 (0.74)	44.44 (0.70)	42.95 (0.78)
χ^2 Caucasian vs Minority	153.48***	138.93***	88.28***	145.25***
Subsidized lunches				
Number	Nm = 2,096	Nm = 2,161	Nm = 2,111	Nm = 2,161
	50.34 (0.65)	50.09 (0.68)	51.70 (0.64)	50.55 (0.72)
Yes	Nm = 2,466	Nm = 2,727	Nm = 2,499	Nm = 2,730
	44.46 (0.60)	43.73 (0.63)	44.59 (0.59)	45.39 (0.67)
χ^2 Subsidy vs No subsidy	94.95***	108.82***	123.95***	61.42***
Age				
	Nt = 4,562	Nt = 4,888	Nt = 4,610	Nt = 4,891
8 years	47.76 (0.52)	46.86 (0.49)	48.79 (0.53)	49.58 (0.52)
10 years	47.40 (0.55)	46.91 (0.58)	48.15 (0.52)	47.97 (0.61)
12 years	47.04 (0.86)	46.96 (0.91)	47.50 (0.81)	46.35 (0.95)
χ^2	0.62	0.01	2.15	12.20***

Nc = Number of cases for between subjects variables. Nm = Number of measurements for each level of Reduced or Free Lunches, which is a discrete, time-varying covariate. Nt = Number of measurements used in computing the effect on each Test of Age, which is a continuous, time-varying covariate. Marginal means for Age are for representative ages.

* p<0.05, ** p<0.01, *** p<0.001.

test were computed for each program participation condition (see Table 2). Note that the three program participation factors together define five program participation conditions: (1) a Matched Controls group of students with no Overall program participation, (2) an Own Controls condition of students with Overall but neither current year nor previous year program participation, (3) a Current Only condition of students with current year but not in the previous year program participation, (4) a Previous Only condition of students with previous year but no current year program participation, and (5) a Current and Previous Year condition of students with both current year and previous year program participation. Table 2 also shows which condition marginal means differ at p<0.05 for each pairwise comparison of marginal means. Because there were ten possible pairwise comparisons of marginal means for each achievement test, the comparison-wise error rate for each test was controlled by using a sequential Bonferroni adjustment to the p-levels for each comparison [45].

Table 2 shows that the same ordering of means across program participation conditions was found for all four achievement tests, although the significance levels varied some-

what across tests. The highest mean achievement test scores were found in the Current and Previous Year condition, followed by the current-only condition, the matched-controls condition, and the own-controls condition. The lowest mean achievement test scores were found in the previous-only condition. For the basic battery test, there were significant effects for overall participation, current-year participation, previous-year participation, and the current-year x previous-year interaction. Pairwise comparisons showed that each condition's marginal mean differed significantly from all other marginal means (all adjusted p's<0.05). For the Math test, again all program participation effects and all pairwise comparisons of condition marginal means were significant.

For the Reading test, the Overall effect of participation was nonsignificant, but significant effects were found for Current Year participation, Previous Year participation, and the Current Year x Previous Year interaction. All pairwise comparisons of condition marginal mean were significant (with adjusted p's<0.05), except for the comparison of the Matched Controls with Own Controls. For the Language test, overall participation and Previous Year participation

Table 2. Test score marginal means (and standard errors) for each program participation condition.

Condition	Test			
	Basic Battery	Reading	Language	Math
Controls, N = 1,100	47.48 _a (0.69)	46.70 _a (0.70)	48.05 _a (0.65)	48.28 _a (0.73)
Program Participants N = 1,106				
Own Controls ^a	44.98 _b (0.74)	45.17 _{bc} (0.77)	46.50 _{bc} (0.74)	45.81 _b (0.82)
	Nm=854	Nm=857	Nm=782	Nm=779
Current Year Only ^b	49.52 _c (0.71)	49.37 _c (0.74)	49.77 _{bc} (0.70)	50.30 _c (0.78)
	Nm=1,179	Nm=1,179	Nm=1,181	Nm=1,170
Previous Year Only ^c	42.54 _d (0.99)	42.34 _d (1.06)	45.67 _{cd} (1.09)	38.74 _d (1.16)
	Nm=197	Nm=197	Nm=197	Nm=195
Current & Previous Year ^d	54.50 _e (1.12)	53.59 _e (1.21)	53.98 _e (1.23)	53.92 _e (1.31)
	Nm=203	Nm=205	Nm=205	Nm=203
Significance of Effect (df=1)	X ²	X ²	X ²	X ²
Overall	7.84**	3.08	3.12	7.19**
Current year	83.80***	58.80***	29.97***	54.51***
Previous year	8.98**	9.75**	0.72	48.87***
Current × Previous	37.71***	27.29***	12.17***	50.83***

Marginal means within a column that do not share common subscripts differ significantly at $p < 0.05$ using Holm's (Holm, 1979) sequential Bonferroni test. Standard errors are in parentheses. N_c = Number of cases for between subjects variables. N_m = Number of measurements for each Program Participation condition.

- ^a – Not in participating schools in current or past year;
 - ^b – Attending a participating school in the current year but not previous year;
 - ^c – Attended a participating school in the previous year but not currently attending;
 - ^d – Attending a participating school currently and had attended one in the previous year.
- * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3. Comparison of means of students who moved from a participating to a non-participating school with means of students who moved from one non-participating school to another.

Test	Test Score Marginal Means (SE)			Test of Significance	
	Previous only ^a (n=197)	Changed Schools controls ^b (n=197)	Primary control group (n=1,100)	Previous only vs. Changed schools (X ²)	Primary control vs. Changed school (X ²)
Basic Battery	42.54 (0.99)	49.97 (2.18)	47.48 (0.69)	11.61***	1.30
Reading	42.34 (1.06)	47.64 (1.92)	46.70 (0.70)	7.62**	0.24
Language	45.67 (1.09)	49.09 (2.5)	48.05 (0.65)	1.87	0.17
Math	38.74 (1.16)	47.90 (1.70)	48.28 (0.73)	29.03***	0.05

- ^a – Attended a participating school in the previous year but not currently attending;
 - ^b – Test scores if student changed schools in the previous year (students matched for demographic variables and year of change with Previous only group).
- * $p < .05$, ** $p < .01$, *** $p < .001$.

were nonsignificant, while Current Year, and the Current Year × Previous Year interaction were significant. All pairwise comparisons of the Current and Previous Year condition marginal mean with the four other condition means were significant. However, the Matched Controls condition

mean did not differ significantly from the Own Controls condition, the Current Only condition, or the Previous Only condition. Still, as with all three other achievement tests, the comparisons of the Current Only condition mean with both the Own Controls and Previous Only condition

means were significant, but unlike all three other achievement tests, the comparison of Own Controls with Previous Only was nonsignificant.

Differential program effectiveness for different demographic groups

We conducted supplemental analyses to test for differential program effectiveness for students in different demographic groups. These analyses included the same effects as in the previous analyses. In addition, they included all two-way interaction effects of Current Year and Previous Year with Gender, Ethnicity, and Subsidized Lunches, as well as their three-way interaction effects. Only 2 of the 36 interactions tested were significant: Current Year \times Previous Year \times Subsidized Lunches on Basic Battery scores, ($X^2=4.75$, $df=1$, $p<0.05$), and Previous Year \times Gender on Math scores, ($X^2=6.26$, $df=1$, $p<0.05$). Given these near chance findings, the program's apparent effects on achievement do not appear to be conditional on gender, ethnicity, or SES.

Effects of changing schools for non-program participants

The decreased achievement in the Previous Only condition was not anticipated, but one obvious potential explanation for this unexpected finding is that changing schools from any school to another school can be disruptive, causing achievement to deteriorate. To test this explanation, a second matched comparison group was constructed. For each of the 197 program students in the Previous Only condition, the first year that the student changed from a program to a non-program school was recorded. For these students, a matched control student was randomly selected based on demographic characteristics (i.e., gender, ethnicity, age, SES), and the first year that the student changed schools. Matched students were found for 190 of the 197 program students using this procedure. For the remaining 7 program students, a control student was randomly selected from the non-program district students who matched the program student on all the primary criteria, and who changed schools in some year from 1996 through 2000.

For each of the 197 control students a Changed Schools variable was created for each year in which test scores were available. This variable was coded as 1 if test scores were available and the school for the previous year differed from the school for the current year. Analyses included main effects for Year, Changed Schools, Gender, Ethnicity, Subsidized Lunches, and the linear effect of Age. Table 3 shows the marginal means (and standard errors) for the years in which students changed schools. The table also displays χ^2 s comparing CAPSLE and control students the year following a school change. All scores for the control students were significantly higher than the scores of those who moved out of a program school ($ps<0.001$), except for the language scores where the decline in the previously-only condition was less pronounced following departure from a CAPSLE school. The secondary and primary control groups' marginal group means on the remaining tests did not differ significantly. These findings indicate that the performance of the matched control group did not deteriorate after changing schools and that the performance drop associated with CAPSLE students who changed schools is most likely associated with changes in the quality of the school environment.

That is, changing from a more protective non-violent environment of the CAPSLE schools to environments that involved no special effort to protect conditions of learning.

DISCUSSION

Children who attended a school participating in the CAPSLE program for two consecutive years performed better on standardized achievement tests than did a carefully matched cohort of children in the same school district who had no exposure to CAPSLE. The proposition that this finding is a result of attendance in a CAPSLE program school is bolstered by within-subjects comparisons showing that children who attended one of these schools for two consecutive years showed an average gain of 8 to 10 percentile points relative to their pre-CAPSLE test scores. These findings could not be accounted for by demographic factors, including age, gender, ethnicity, and low income or coincidental changes associated with a particular time period when the program was introduced. The use of a multiple baseline design further strengthens the claim that improvements in academic performance appear to coincide with the school implementing CAPSLE.

This study, to our knowledge, is one of the first of a school-wide violence prevention program to examine educational attainment as an outcome. Potential mechanisms for producing achievement gains appear to center on increased engagement in instructional time and decreased distractions from disruptive behavior as well as decreased concerns regarding conflict. Although these data do not provide direct support of this conclusion, a previous, more detailed, study of one of the schools reported here demonstrated decreases in disciplinary referrals, school exclusions, and an increased experience of safety on the part of students associated with the introduction of the program [39]. Improvement in academic performance is all the more impressive as CAPSLE is a purely social systems oriented intervention without any academic components that are likely to have a direct impact on language, mathematics, or reading skills of the students. In fact, the implementation of the program requires a modest amount of classroom time for class discussions and counselor-led instruction about power struggles and conflict. Previously reported research on the CAPSLE project supports an interpretation of educational benefits of CAPSLE as derivable through improved teaching opportunities for teachers in CAPSLE schools [39].

The current results offer no support for the view that more advantaged children (i.e., higher income, non-minority) experienced the greatest gains in academic achievement through attendance in CAPSLE schools. We observed only two, marginally significant interactions that pointed in the direction of greater benefit for non-minority and no lunch subsidy groups, but in view of the number of interactions tested, we can have no confidence in these observations. In general the results indicate that children showed notable progress, regardless of economic and cultural factors, when given the opportunities provided by a peaceful school atmosphere.

Notably, the children who left a CAPSLE school showed signs of a decline in academic achievement, both relative to their performance within the CAPSLE school and relative to

the performance of students who were not in the CAPSLE program. This deterioration was not attributable simply to changing schools (e.g., the need to adjust to a new school environment or the selective departure of students currently experiencing educational problems from schools). If the changes in schools were mostly triggered by family mobility in both past CAPSLE school participants and matched control students, the relative poor performance of children in the Previous Only group is akin to a treatment withdrawal condition in an A-B-A design [46]. A decline toward baseline following the withdrawal of treatment would normally be considered as strengthening the assumption of a causal relationship between previous improvements and the intervention conditions. However, in the present instance the decline appeared to go below baseline both relative to Own Control and Matched Control levels of attainment. This finding suggests a detrimental effect on educational attainment when a child shifts from a school environment in which bully-victim-bystander problems are addressed in a systematic and reflective manner to a more typical school setting where a variety of responses to these problems exist, including ignoring all but severe incidents.

Limitations

Despite these intriguing and compelling results, this study has several important limitations. First, the first participating school in the study was not randomly chosen but rather started at a time of crisis. Thus, regression to the mean would be expected. However, schools recruited to the study later were selected from schools volunteering for the program. The second school to join had been a comparison school with a number of behavioral measures already in place but the start of the program was not associated with any particular incident. The remaining schools to join were enrolled as part of a school district wide study. It is therefore unlikely that regression to the mean could account for the observations reported here.

Second, the CAPSLE schools received the special attention usually associated with new initiatives. It remains to be seen whether the effects noted here can be attributed specifically to CAPSLE, as opposed to the introduction of any school-wide program, or perhaps could be produced by any radical changes in curricula or instructional approaches. CAPSLE was the only new program introduced to the school district over the study period. Although the introduction of new programs (except statutory programs) was deferred until the end of the study period in CAPSLE schools, children in the Matched control group came from schools in the district with a variety of fresh initiatives. In fact, we intended to model this impact in order to strengthen the control condition; however, the sheer number and diversity of patterns of implementation across the school district made this impossible with our current sample size. Nevertheless, the absence of an alternative intervention condition and the lack of random assignment are important limitations of the study. Comparing CAPSLE with alternative interventions using random assignment to conditions is clearly needed.

Third, the intervention in the first school was did not have a complete manual, and the intervention protocol was refined and modified as information emerged. Although a manual to facilitate replication was completed during the course of

the study, no fidelity checks were done to ensure that the manual was implemented with reasonable fidelity.

Finally, the program has multiple components and at this stage, we do not know if all of them are essential. It would be risky to generalize from the results of a small number of intervention schools. To establish the program as effective it needs to be implemented in more schools simultaneously. Replications are in progress with random assignment of schools to program.

Educational implications

Previous studies highlight the importance of peer modeling in school conflict mediation and academic achievement [47,48]. Our program, which did not incorporate parents, supports these findings. Changing the peer role models for the children may be more important than complex, expensive, and often unsuccessful interventions with parents and individual families. Specifically, changing the peer relationships appears to change the social norms within the schools. Such changes to an overall school climate are not only conducive to academic achievement but also serves the socialization mission of the schools.

CONCLUSIONS

The CAPSLE program is cost effective because it makes use of resources that are typically readily available to schools. The greatest expense is in initial investment of time of school staff, but this investment pays significant dividends in terms of quality educational time brought about the improvements in classroom climate. Whether such a program is useful in primary prevention (i.e., in schools without crises) remains to be seen. In short, the CAPSLE program is cost effective, non-pathologizing, simple in design, easy to implement, and supportive of the educational atmosphere of schools.

Acknowledgements

Data for this study were collected as part of the Peaceful Schools Project directed by Drs. Fonagy, Twenlow, and Vernberg, with special acknowledgment to the Jack Taylor Family Foundation. We gratefully acknowledge the contributions of Anne Jacobs, Michael Wright, Larry Wolgast, and Patricia McElliot in carrying out the project and thank the students and staff of the Topeka, Kansas public 504 school district for their participation in this research.

The authors wish to express their deepest gratitude and affection for the late Jim Fultz PhD, who made tremendous personal and scientific contributions to the Peaceful Schools Project.

REFERENCES:

1. Farrell AB, Meyer AL, Emig EM, Sullivan TN. Development and evaluation of school-based violence prevention programs. *Journal of Clinical Child Psychology*. 2001; 30: 207-20.
2. Embry DD, Hancock DJ, Vazsonyi M, et al. Peacebuilders, a theoretically driven, school-based model for early violence prevention. *American Journal of Preventive Medicine*. 1996; 12: 91-100.
3. Gabriel KM, Hopson L, Haskins M, Powell KE. Building relationships and resilience in the prevention of youth violence. *American Journal of Preventive Medicine*. 1996; 12: 35-37.