Creating a Peaceful School Learning Environment: The Impact of a Violence Prevention Program on Educational Attainment

Peter Fonagy
University College of London, England
Email p.fonagy@ucl.ac.uk
Tel. 020 7679 1791
Fax 020 7916 8502

Stuart W. Twemlow
Erik Erikson Institute for Research & Education
Stockbridge, Massachusetts

Eric Vernberg
University of Kansas at Lawrence
Jim Fultz
The Menninger Clinic
Topeka, Kansas

Frank C. Sacco
Community Services Institute
Springfield, Massachusetts
Abstract
The impact of a violence prevention program on the school learning climate in five elementary schools over a 5-year period was investigated using a multiple baseline design. Academic attainment test scores of 1,106 students were monitored before and after the introduction of the program across the school district and were contrasted with the attainment scores over the same period of 1,100 individually matched students from the school district who were not in schools that joined the program. Program participation was associated with highly significant improvements in individual achievement test scores as well as notable reductions in scores in those students who left schools with active programs. A simple, low cost anti violence intervention, designed to reduce peer conflicts that interfere with the educational process, appears to significantly benefit educational performance of children in participating elementary schools.
Creating a Peaceful School Learning Environment: The Impact of a Violence Prevention Program on Educational Attainment

Violence prevention has become an increasingly important activity in the United States and there are powerful rationales for basing these efforts in school settings (Farrell, Meyer, Kung & Sullivan, 2001). There is little doubt that violence among peers is problematic in many North American schools (Embry, Flannery, Vazsonyi, Powell, & Atha, 1996; Gabriel, Hopson, Haskins, & Powell, 1996; National Center for Injury Prevention and Control, 1999), and there appears to be a propensity for schools to generate informal social norms that predispose to violence (Fagan & Wilkinson, 1998; Garbarino, 1995; Wiist, Jackson, & Jackson, 1996). School-based programs attempt to challenge and change these norms and reduce the level of overt and relational aggression among peers. The degree of violence among peers has also been demonstrated to be a key mediator of the relation between family factors and engagement in acts of serious delinquency or violence (e.g., Henry, Tolan & Gorman-Smith, 2001).

Elementary schools may be particularly desirable settings for prevention. Developmental studies repeatedly emphasize that the early onset of aggression and violence predicts a child’s future risk for violence (Flannery & Williams, 1999). Children’s beliefs about aggression and maladaptive attributional biases concerning the hostile intent of others tend to crystallize between the ages of 6 and 12 (Aber, Brown, Chaudry, Jones & Samples, 1996). Middle childhood is an important time for the integration of emotion regulation and cognitive and behavioral skills needed to develop social competence (Greenberg, Craighead, Evans, & Craighead, 1995). Elementary schools also provide relatively ready access to high risk samples and relevant outcome data (Eddy, Dishion, & Stoolmiller, 1998; Samples & Aber, 1998).

There are a number of reasons why educational attainment merits consideration as an important outcome of violence prevention programs. Violence among elementary school-age
children, in the form of overt and relational aggression, potentially interferes with the learning process in many ways. Children who are repeatedly targeted as victims of aggression show elevated symptoms of depression and anxiety, which may interfere with concentration and task completion (Vernberg, 1990; Vernberg, Abwender, Ewell, & Beery, 1992). Anger, embarrassment, or fear related to aggressive acts among peers directs children’s energies toward revenge or avoidance rather than academic work, and ongoing conflicts among children disrupt instructional time. Among children who perpetrate violence towards peers, educational problems are common. For example, highly aggressive youth are more likely than less aggressive peers to have been retained in one or more grades (Rodney, Crafter, Rodney, & Mupier, 1999), and the combination of poor academic achievement and early conduct problems has been repeatedly demonstrated to predict criminal involvement across the lifespan (Farrington, 1994; Lewis, Yeager, Lovely, Stein, & Cobham-Portorreal, 1994; Patterson, DeBarysh, & Ramsey, 1989; Loeber & Farrington, 2000). There is a well established finding which links behavioral problems to reading and language difficulties (Beitchman, Brownlie, Inglis et al., 1996; Beitchman, Brownlie, & Wilson, 1996; Rutter et al., 1997), with more recent studies suggesting that the association of language problems with behavioral disorders are mediated by reading difficulties (Tomblin, Zhang, Buckwalter, & Catts, 2000). Conversely, educational attainment has been shown to be a protective factor, an indicator of so-called resilient children (Quinton, Pickles, Maughan, & Rutter, 1993; Rutter, 1985; Rutter, Giller, & Hagell, 1998).

Enhanced educational attainment is a concern, not just for individuals involved in violence as aggressors or as victims, but for the entire community. School violence that adversely affects the educational process depletes the value of the human capital of a town, of a
state or of a nation. There are well-known 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 (Rutter, Maughan, Mortimore, J., Ouston, & Smith, 1979). Educational attainment is a desirable outcome measure because of its strong psychometric properties (Newman & Ciarlo, 1994) that include sensitivity to change, resistance to reactivity, various forms of response bias and 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 school boards, school administrators, principals and teachers by creating a common cause between mental health professionals and educators (Elias et al., 1997; Vernberg & Biggs, in press).

To our knowledge, however, there have been no large scale studies that examine the impact of elementary school based violence prevention programs on children’s educational attainment. A number of studies have demonstrated promising effects for violence prevention programs in elementary schools, primarily in the realms of 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, the teaching of non-violent alternatives to conflicts, the enhancement of negotiating, and 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 in Topeka, Kansas, which started with an intervention in a single elementary school in 1994 (Twemlow, Fonagy et al., 2001a), is an open social systems-psychodynamically informed intervention
integrating several approaches that address 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 (Twemlow, 2000; Twemlow, Fonagy, & Sacco, 2001b). 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 program addresses children as well as school employees, volunteers, and parents of children at the school. Schools with high levels of bully-victim problems usually have greater teacher dissatisfaction and turnover, and low levels of parental involvement. High out-of-school suspension rates and low levels of academic achievement are also common. Such schools often lack a plan to address the overall school social climate, and rely heavily on punitive or coercive discipline strategies.

The CAPSLE program utilizes four primary components. A **positive climate campaign** uses class discussion, counselor-led lessons, posters, magnets, bookmarks, and other devices to encourage a shift in language (and thinking) of all students and personnel. These language tools allow the identification and resolution of problems that occur when coercive power dynamics dominate the school environment (Twemlow et al., 2001b). A **classroom management plan** assists teachers to discipline by focusing on the understanding and correction of problems rather than on punishment and criticism. A **physical education program** derived from a combination of role-playing, relaxation techniques, and defensive Martial Arts teaches children skills to handle victimization and bystanding behavior. This program helps children protect themselves and others with nonaggressive physical and cognitive strategies. Schools also put in place one of two support programs: **peer mentorship** or **adult mentorship**. These relationships provided additional
containment and modeling to assist children in mastering the skills and language to deal with power struggles.

During the period 1996 - 2000 the program extended out to 5 schools in the school district. The gradual rollout of the program enabled us to contrast the performance of children who attended schools that participated in the program with those who were in other schools in the same school 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 prior to the school joining the program, their performance prior to their individually joining a participating school or when leaving the school to move to another school where the program was not available. The study aimed to identify whether improvements in educational attainment were associated with attendance in a school where the program was implemented.

Method

Program Implementation

The CAPSLE program was operating in two public schools in the Topeka, Kansas, school district by 1996 (see Twemlow, Fonagy et al., 2001a; Twemlow, Fonagy, Sacco, Gies et al., 2001 for description of program development and initial evaluation). The program was implemented in three additional schools in the same district at the beginning of the 1999-2000 academic year. The four intervention components described above were implemented by school staff 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 Stephen Twemlow, a martial arts specialists. School counselors or social workers coordinated and supervised the peer mentor or adult mentor program. All
Violence prevention interventions were described in a detailed manual designed to enhance program fidelity and replicability (Twemlow, Sacco, & Twemlow, 1999).

Participants

Standardized achievement test scores were available for 1106 students who attended a CAPSLE program school during a school year in which the program was implemented for one or more years from 1996 through 2000. The first year for which achievement scores were available was 1996 for 156 students (14.1%), 1997 for 198 students (17.9%), 1998 for 236 students (21.3%), 1999 for 212 students (19.1%), and 2000 for 304 students (27.5%). There were 587 male students (53.1%) and 519 female students (46.9). The majority of students were white (58.4%). For the first year for which achievement scores were available, the majority of the students received free or reduced-price lunches (59.4%), and ages ranged from 7 to 14 years, with a mean age of 9.7.

Matched Control Group

To provide a comparison of achievement test scores for program participants with test scores for other students in the district, a matched comparison group of students was constructed. First, for each student who participated in the CAPSLE program, the student’s gender and ethnicity were recorded. Then, for the first year for which achievement test scores were available for the student, the year and the student’s age were recorded, as well as whether free or reduced-price lunches were received during the school year. Finally, a matched control student was randomly chosen from the district students who did not attend a CAPSLE program school in any year from 1996 through 2000, and whose gender, ethnicity, age, receipt of free or reduced-price lunches, and first year for which test scores were available matched the program student’s characteristics. Using this procedure, a total sample of 2206 students was achieved, with 1106
program students and 1100 control students (complete demographic information was not available for six of the program participants, so no matched controls could be chosen for these students).

**Achievement Test Scores Prior to Program Implementation and After Transferring to a non-Program School**

In addition to the comparison of achievement test scores provided by the matched control group, it also was desirable to obtain a comparison of achievement test scores prior to and after program implementation for program participants. 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.

[here would be a good place to describe the number of program students with test scores from 1996-2000 who transferred to a non-program school each year]

**Measures of Student Achievement**

Academic achievement was assessed using the Metropolitan Achievement Test (MAT) 1993. This test 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 scores are standardized to national norms and individual scores are reported as percentiles. In this analysis, we looked at composite test scores (Basic Battery) and Reading, Language, and Math test scores to measure student achievement.
Demographic Group Differences on Achievement Test Scores

Preliminary analyses were performed to assess differences on achievement test scores across demographic groups. For each student in the total sample, the student's group membership was coded for a Gender variable (0 = male; 1 = female) and an Ethnicity variable (0 = white; 1 = non-White). For years for which test scores were available, receipt of Free or Reduced-Price Lunches was coded for each year (0 = not received; 1 = received). The student's Age (in years) also was recorded for years for which test scores were available.

Students’ achievement test scores were analyzed using the BMDP 5V program for unbalanced repeated measures models with structured covariance matrices (Dixon, 1992). Year was treated as the within-subject, repeated measures factor. The within-subject design was unbalanced because there were different numbers of years and different 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. In the preliminary analyses and all subsequent analyses, the estimated correlations between students’ test scores across Years were uniformly high (ranging from 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.

In the preliminary analyses, Gender and Ethnicity were treated as fixed covariate factors, receipt of Free or Reduced-Price Lunches was treated as a time-varying covariate factor, and Age was treated as a time-varying continuous covariate. Separate analyses were performed for each of the four achievement test scores. For each achievement test, the analysis provided
estimates of the effects of each demographic variable which were used to compute estimated test score marginal means for each group on that variable, controlling for the influences of the other demographic variables. Table 1 shows these marginal means for the Basic Battery, Reading, Language, and Math achievement test scores.

As can be seen in Table 1, with a few exceptions there is a quite consistent pattern of differences in demographic group marginal means across all four achievement tests. Females’ scores were significantly higher than males’ scores on Basic Battery ($M'$s = 48.30 and 46.50, respectively, $X^2 = 4.80, df = 1, p < .05$), on Reading ($M'$s = 47.93 and 45.89, respectively, $X^2 = 6.70, df = 1, p < .01$), and on Language ($M'$s = 50.10 and 46.19, respectively, $X^2 = 25.72, df = 1, p < .001$), but females’ scores were nonsignificantly lower than males’ scores on Math ($M'$s = 47.46 and 48.48, respectively, $X^2 = 1.53, df = 1, p > .2$).

White students’ scores were significantly higher than non-white students’ scores on Basic Battery ($M'$s = 52.51 and 42.29, respectively), on Reading ($M'$s = 51.58 and 42.24, respectively), on Language ($M'$s = 51.85 and 44.44, respectively), and on Math ($M'$s = 52.98 and 42.95, respectively), all $X^2$s > 88.28, df's = 1, p's < .001.

Also, scores for students who did not receive Free or Reduced Lunches were significantly higher than scores for students who received Free or Reduced Lunches on Basic Battery ($M'$s = 50.34 and 44.46, respectively), on Reading ($M'$s = 50.09 and 43.73, respectively), on Language ($M'$s = 51.70 and 44.59, respectively), and on Math ($M'$s = 50.55 and 45.39), all $X^2$s > 61.41, df's = 1, p's < .001.
There were no significant effects for Age on Basic Battery, Reading, and Language (all p’s > .1). On the Math test, however, there was a significant effect for Age such that younger students had higher scores, $b = -.81, \chi^2 = 12.20, df = 1, p < .001$.

The preliminary analyses showed that there are marked differences in test scores across several demographic groups, so this made it important to control for such differences in testing the effects of participation in the CAPSLE program on student achievement.

Results
The effects of a total of eight factors on students' Basic Battery, Reading, Language, and Math achievement test scores 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) Year, (2) Overall participation, (3) Current Year participation, (4) Previous Year participation, (5) Gender, (6) Ethnicity, (7) receipt of Free or Reduced-Price Lunches, and (8) Age.

Depending on students' age, residency in the school district, and the achievement-testing schedule, test scores were available for one or more Year from 1993 through 2000. If a student attended a Peaceful School program school during a school year in which the program was implemented for any year from 1996 through 2000, this defined Overall participation in the program. For years for which test scores were available for students who participated in the program, Current Year participation and Previous Year participation in the CAPSLE program were recorded. For all three of these participation variables, nonparticipation was coded as a 0 and participation was coded as a 1.

Achievement test scores were analyzed using as predictors the same demographic variables used in the preliminary analyses, and additionally including Overall Participation as a
fixed covariate factor, and Current Year participation and Previous Year participation as time-varying covariate factors. Thus, the design for testing the effectiveness of CAPSLE program participation in facilitating achievement included main effects for Year, Overall participation, Current Year participation (nested within Overall participation), and Previous Year participation (nested within Overall participation), the Current Year x Previous Year interaction effect (nested within Overall participation), main effects for Gender, Ethnicity, and receipt of Free or Reduced-Price Lunches, and the linear effect of Age.

Effects of Program Participation on Achievement Test Scores

Separate analyses were performed for each of the four achievement tests. Estimated test score marginal means for each test were computed for each program participation condition, controlling for influences of the demographic variables on test scores. 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 no 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 shows these marginal means for the Basic Battery, Reading, Language, and Math achievement test scores, respectively, for each program participation condition. Additionally, Table 2 also shows which condition marginal means differ at $p < .05$ 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 Holm’s (Holm, 1979) sequential Bonferroni adjustment to the p-levels for each comparison.

---

Inspection of the means in Table 2 shows that exactly the same ordering of means across program participation conditions was found on all four achievement tests, although the significance of differences between means for specific pairs of conditions varied somewhat across tests. The highest mean achievement test scores were found in the Current and Previous Year condition, followed in turn by the Current Only Condition, the Matched Controls condition, and the Own Controls condition, and 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 ($X^2 = 7.84, df = 1, p < .01$), Current Year participation ($X^2 = 83.80, df = 1, p < .001$), Previous Year participation ($X^2 = 8.98, df = 1, p < .01$), and the Current Year x Previous Year interaction ($X^2 = 37.71, df = 1, p < .001$). Pairwise comparisons showed that each program participation condition marginal mean differed significantly from all other marginal means (all adjusted p’s < .05).

Similarly, for the Math test, there were significant effects for Overall participation ($X^2 = 7.19, df = 1, p < .01$), Current Year participation ($X^2 = 54.51, df = 1, p < .001$), Previous Year participation ($X^2 = 48.87, df = 1, p < .01$), and the Current Year x Previous Year interaction ($X^2 = 50.83, df = 1, p < .001$). Again, all pairwise comparisons of condition marginal means were significant (all adjusted p’s < .05).

For the Reading test, there was a nonsignificant effect for Overall participation ($X^2 = 3.08, df = 1, p > .07$), but significant effects for Current Year participation ($X^2 = 58.80, df = 1, p$...
Violence prevention

< .001), Previous Year participation ($X^2 = 9.75, df = 1, p < .01$), and the Current Year x Previous Year interaction ($X^2 = 27.29, df = 1, p < .001$). All pairwise comparisons of condition marginal mean were significant (with adjusted $p$’s < .05), except for the comparison of the Matched Controls ($M = 46.70$) with Own Controls ($M = 45.17$), $X^2 = 3.08, df = 1, adjusted p > .07$. For the Language test, there were nonsignificant effects for Overall participation ($X^2 = 3.12, df = 1, p > .07$) and Previous Year participation ($X^2 = 12.17, df = 1, p < .001$). All pairwise comparisons of the Current and Previous Year condition marginal mean ($M = 53.98$) with the four other condition means were significant (with adjusted $p$’s < .001). However, the Matched Controls condition mean ($M = 48.05$) did not differ significantly from the Own Controls condition mean ($M = 46.50$), $X^2 = 3.12, df = 1, adjusted p > .2$, the Current Only condition mean ($M = 49.77$), $X^2 = 4.62, df = 1, adjusted p > .1$, or the Previous Only condition mean ($M = 45.67$), $X^2 = 4.25, df = 1, adjusted p > .1$. 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 (with adjusted $p$’s < .001), but unlike all three other achievement tests, the comparison of Own Controls with Previous Only was nonsignificant, ($X^2 < 1.0, ns$).

In addition to the effects of program participation, there were a number of significant effects of the demographic variables on achievement test scores. The patterns of effects for these variables were virtually identical to the patterns found in the preliminary analyses involving only the demographic variables. Females scored significantly higher than males on the Basic Battery, Reading, and Language tests (all $p$’s < .05), but on the Math test females scored nonsignificantly lower than males ($p > .2$). On all four tests, white students scored significantly higher than non-
white students (all $p$’s < .001), and students who did not receive Free or Reduced-Price Lunches scored higher than students who did receive Free or Reduced-Price Lunches (all $p$’s < .001).

Age was not significantly related to Basic Battery, Reading, or Language test scores (all $p$’s > .2). The only effect that differed from what was found in the preliminary analyses was for Age on Math test scores. This effect was nonsignificant, $b = -.44$, $X^2 = 3.59$, $df = 1$, $p > .05$, whereas the negative relationship of Age with Math scores was significant in the preliminary analyses.

The effects of the demographic variables were controlled for in computing the marginal means for the program participation conditions presented in Table 2, but it also was of interest to determine if there was evidence of differential program effectiveness for students in different demographic groups. Accordingly, follow-up analyses were performed to explore this possibility.

**Differential Program Effectiveness for Different Demographic Groups**

The design of the follow-up analyses included the same effects as in the previous analyses, and additionally included the two-way interaction effects of the Current Year participation factor with Gender, Ethnicity, and receipt of Free or Reduced-Price Lunches, the two-way interaction effects of the Previous Year participation factor with Gender, Ethnicity, and receipt of Free or Reduced-Price Lunches, and the three-way interaction effects of the Current Year participation factor and the Previous Year participation factor with Gender, Ethnicity, and receipt of Free or Reduced-Price Lunches.

Across the four achievement tests there were a total of 36 participation factor by demographic factor interactions that were tested, and only two of these 36 interactions were found to be significant. There was a significant Current Year x Previous Year x receipt of Free or Reduced-Price Lunches interaction on Basic Battery test scores, $X^2 = 4.75$, $df = 1$, $p > .05$, and
a significant Previous Year participation x Gender interaction on Math test scores, $X^2 = 6.26$, df = 1, $p > .05$. To simplify interpretation of these two significant interactions, simplified versions of the Basic Battery and Math analyses were rerun. The analysis of Math scores was rerun dropping all nonsignificant participation factor by demographic factor interaction effects. The Basic Battery analysis was rerun dropping all nonsignificant participation factor by demographic factor interaction effects, except for the Current Year participation x receipt of Free or Reduced-Price Lunches interaction and the Previous Year participation x receipt of Free or Reduced-Price Lunches interaction, both of which were contained in the significant three-way interaction involving receipt of Free or Reduced-Price Lunches. Marginal means were computed after rerunning the analyses.

For Basic Battery scores, the means for students who did not receive Free or Reduced-Price Lunches first were examined. Compared with no Current or Previous participation ($M = 47.73$), there is a nonsignificant decrease in scores for Previous Only participation ($M = 44.87$), $X^2 = 6.56$, df = 1, adjusted $p > .08$, but there is a increase in scores as a result of Current Only participation ($M = 52.65$), $X^2 = 50.84$, df = 1, adjusted $p < .001$. Compared with Current Only participation, there is a further increase as a result of Current and Previous participation ($M = 60.21$), $X^2 = 82.14$, df = 1, adjusted $p < .001$. For students who did receive Free or Reduced-Price Lunches a similar pattern was observed. However, the increase in Basic Battery test scores resulting from Current and Previous participation was not statistically significant. Compared with no Current or Previous participation ($M = 42.15$), the decrease in scores for Previous Only participation ($M = 40.18$) was not statistically significant, $X^2 = 3.07$, df = 1, adjusted $p > .6$, but the increase in scores as a result of Current Only participation ($M = 46.43$) remained statistically significant, $X^2 = 46.65$, df = 1, adjusted $p < .001$. Compared with Current Only participation,
there is a further increase as a result of Current and Previous participation ($M = 48.41$), but this increase did not approach significance, $X^2 = 2.20$, ns. In addition to inspection of individual pairwise comparisons, the range of means across program participation conditions for students who did not receive Free or Reduced-Price Lunches, range = 60.21 - 44.87 or 15.34, can be compared to the range for students who did receive Free or Reduced-Price Lunches, range = 48.41 - 40.18 or 8.23. The range for students who did receive Free or Reduced-Price Lunches is roughly half as large.

For Math scores, pairwise comparisons showed that for no Previous Year participation, females’ scores ($M = 47.54$) and males’ scores ($M = 48.56$) did not differ, $X^2 = 1.54$, df = 1, ns., and neither of these mean scores differed from the mean for Previous Year participation for females ($M = 47.23$), $X^2's < 1.1$, ns. Previous Year participation scores for males ($M = 45.61$) were significantly lower than the mean for no Previous Year participation for males, $X^2 = 8.31$, df = 1, adjusted $p < .01$, but did not differ from the scores for females with no Previous Year participation or with Previous Year participation, $X^2's < 2.4$, df’s = 1, adjusted p’s > .6.

Effects of Changing Schools for Non-Program Participants

The decreased achievement found 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 secondary matched comparison group was constructed. For each of the 200 program students who contributed test scores to the Previous Only condition, the first year from 1996 through 2000 in which the student changed schools from a program school to a non-program school was recorded. Then, for each of these students, a matched control student was randomly selected on the same criteria that were used for constructing the primary matched
comparison group (i.e., gender, ethnicity, age, free or reduced-price lunch status), and whose first year from 1996 through 2000 in which the student changed schools matched the first year in which the program student changed schools from a program to a non-program school. Non-program matched control students were found for 190 of the 200 program students using this procedure. For the remaining 10 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 200 control students in the secondary matched controls sample, a Changed Schools variable was coded for each year from 1993 through 2000 in which test scores were available. Changed Schools was initially coded as 0 for each year, and then recoded as 1 for years in which the test scores were available and the school for the previous year differed from the school for the current year. Achievement test scores for the matched controls were analyzed using a design that included main effects for Year, Changed Schools, Gender, Ethnicity, receipt of Free or Reduced-Price Lunches, and the linear effect of Age. For years in which the control students changed schools, marginal means (and standard errors) for the Basic Battery, Reading, Language, and Math achievement tests, were 49.97 (2.18), 47.64 (1.92), 49.09 (2.5), and 47.90 (1.70), respectively, which were significantly higher than the corresponding Basic Battery, Reading, and Math marginal means in the Previously Only condition ($\chi^2$s > 11.0, df's = 1, p's < .001), and nonsignificantly higher than the corresponding Language marginal mean in the Previously Only condition ($\chi^2$ = 3.14, df = 1, ns). The secondary matched control group Basic Battery marginal mean was significantly higher than the primary matched control group Basic Battery marginal mean ($\chi^2$ = 5.52, df = 1, p < .05), but the secondary and primary control groups marginal group means on the remaining tests did not differ significantly. Thus,
there was no evidence in the secondary matched control group of deterioration of achievement due to changing schools.

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.

This is, to our knowledge, one of the first studies of a school-wide violence prevention program to examine educational attainment as an outcome. Potential mechanisms for producing achievement gains can be articulated, and center on increases in engaged instructional time and decreased distractions from disruptive behavior and concerns regarding conflict. Although the CAPSLE program requires a modest amount of classroom time for class discussions and counselor-led instruction about power struggles and conflict, academic progress apparently is facilitated.

Previously reported research on the CAPSLE project supports this interpretation. In addition to the improvement in academic achievement found here, a study of the first three years of this project found significant reductions in out of school suspensions and other serious infractions involving physical violence together with a reduction of general incidents not involving physical aggression. (Twemlow, Fonagy, Sacco, Gies et al., 2001). The older children appeared to be quickest to benefit from the program, as might be expected from the
developmental age of the children who usually have more comfort independent of parental figures (Twemlow et al., 2001a). The program also seemed to help dependent and withdrawn children especially girls in K-3. They were observed to be less victimized and also to show a reduction in hostility to others with at times troubling (to teachers) and uncharacteristic assertiveness in class. This was not unexpected in view of the fact that the program teaches skills encouraging high levels of physical activity and prompts children to be more outgoing. Reports from schoolteachers supported the fact that many previously passive and withdrawn victimized children often became more verbal and outspoken as the program progressed.

The current results also indicate that more advantaged children (i.e., higher income, non-minority) experienced the greatest gains in academic achievement through attendance in CAPSLE schools. This finding might be expected from common sense inference in that the learning of many of these children is inhibited primarily by disruption in the classroom and on the playgrounds, rather than by both school and family or societal factors. Still, children showed notable progress, regardless of economic and cultural factors, when given the opportunities provided by a peaceful school atmosphere.

It is also noteworthy that children who left a CAPSLE school showed signs of a decline in academic achievement, suggesting either a deterioration or stagnation of academic performance after shifting to a non-CAPSLE school. Additional analyses showed this finding could not be attributed solely to simply having any change of schools. This suggests there may be 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, constructive manner to a more typical school setting where a variety of responses to these problems exist, including ignoring all but severe incidents. A potential explanatory mechanism for this result can be found
in interruption (or discrepancy) theory, which posits that emotional upheaval is created by unexpected discrepancies between social values and norms for behavior following changes in social environments (Mandler, 1990). These changes are often occasioned by relocation, and require shifts in knowledge structures such as schemas, scripts, and other internal representations of social relationships. When these discrepancies result in unpleasant consequences, intense negative affect is believed to occur. If bully-victim-bystander dynamics are highly salient aspects of children’s school environments, a notable shift in norms and behaviors may require a diversion of cognitive resources away from education and towards cognitive accommodation and emotion management.

Previous studies (Powell, Muir, McLain, & Halasyamani, 1996; Sprinthall, Gerler, & Hall, 1992) suggest an important role for peer modeling in school conflict mediation and academic achievement. Our program, which did not work intensively with parents, seems to support these findings. Changing the peer role models for the children maybe more important than complex and expensive and often unsuccessful interventions with parents and individual families. Changing the peer relationships often changes the social norm within the schools and thus the school climate.

This program is cost effective since it makes use of resources usually readily available to schools. The greatest expense is in personnel time of school staff. Whether such a program is useful in primary prevention, that is in schools without crises, remains to be seen. Since it is cost effective, non-pathologizing and simple in design and does not interfere with the educational atmosphere of the school, it should be acceptable to most schools.

Despite these interesting and compelling results, this study has several limitations. In the pilot study, allocation of experimental and control status was not random and the experimental
school started at a time of crisis. Thus, regression to the mean would be expected. However, on most of the important variables there was an improvement that went significantly beyond the baseline levels prior to the crisis in 1994-95 (Twemlow et al., 2001a). Second, the experimental group 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 other violence prevention programs, or perhaps could be produced by changes in curricula or instructional approaches. Certainly, the school system participating in this study implemented numerous attempts to improve educational attainment over the four-year period covered here, although these were distributed equitably across the district. It would be important to compare CAPSLE with alternative interventions, and to use random assignment of intervention conditions. The intervention in the pilot schools was not manualized initially, and the intervention protocol was refined and modified as outcome information emerged. Work on a manual to facilitate replication was completed during the period covered by this study. The program had multiple components and at this stage, we do not know if all these 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. Other replications are in process with randomized assignment of elementary schools to the program.

In light of these promising results, it is worthwhile to describe some of the principles guiding this intervention approach. In previous work (Twemlow, 1995a; 1995b, Twemlow, 2000), we have postulated a dialectical relationship between victims of violent attack and their victimizers. The point of such a dialectic is that the roles are interdependent since neither can exist without the other. The maintenance of these roles is facilitated by the bystanding
“audience” whose interaction with the victim-victimizer dialectic can facilitate more or less victimization. In another context (Twemlow, Sacco, & Williams, 1996), we stated the following: “Schools are stages for the evolution of dramas involving in the interplay of the villains (bullies) and their antagonists (victims) provoked on by the audience (bystanders). Pathological patterns evolve when there are few positive role models in the drama provided by either students or school personnel. Thus, students, teachers, and school personnel, such as the custodian or security staff, can assume any one of the roles in the triad of victim-bully-bystander” (p. 303). This point of view implies that any effective intervention must address the school as a whole rather than individual perpetrators of violent acts.

It was hypothesized that more symmetrical power dynamics should result in reduction in disruption both in the classroom and with fewer competitive struggles on the playgrounds and in school corridors, but the ultimate hope was for a more compassionate and group focused school climate with improved academic performance. The school climate should then become one in which the students and teachers helped each other and served group goals rather than competed with each other striving to achieve dominant status. Further, it is our impression that for programs to be effective multiple components are necessary. The Gentle Warrior program is a psychoeducational method to give school children psychological skills to enable them to deal personally with bullying and their own tendencies to act as a victim or to be bystanders and to observe that behavior and comment on it in others. The Bruno and Peer Mentorship approaches acknowledge the importance of the older individual in acting as a model for containment of the children’s aggressiveness and competitiveness. These models assist children in providing solutions to the personal effects of asymmetrical coercive power relationship on mood and cognition. Thus, all programs have both a psychoeducational and a psychodynamic function.
The Positive Climate Campaign and Disciplinary (Classroom management) Plan assists in changing the language of the various participants in the school climate (including teachers) so that there is a heightened awareness of asymmetrical coercive power relationships and an increase in skills to cope with them. One goal was to change the climate so that social status in highly coercive roles, for example the macho bully, will have less social status than more compassionate and respectful roles. Ideally then, social status is achieved mainly in non-bullying, non-victimized, non-bystanding roles.

These methods appeared to shift these children to a less anxiety provoking, more relational and compassionate mode of functioning. As a result the child becomes more reflective (Fonagy, 1998) and less reactive, with an increased compendium of response options that do not involve bullying, coercion or anxious depressed retreat. Research examining these postulates is critically needed, and evidence on possible educational benefits from a program such as CAPSLE should prove invaluable in convincing school systems and funding sources of the need and promise for such efforts.

Table 1. Test Score Marginal Means for Each Demographic Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Test</th>
<th>Basic Battery</th>
<th>Reading</th>
<th>Language</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Basic Battery</td>
<td>46.50</td>
<td>45.89</td>
<td>46.19</td>
<td>48.48</td>
</tr>
<tr>
<td></td>
<td>N = 1169</td>
<td>(0.68)</td>
<td>(0.70)</td>
<td>(0.64)</td>
<td>(0.73)</td>
</tr>
<tr>
<td>Female</td>
<td>Basic Battery</td>
<td>48.30</td>
<td>47.93</td>
<td>50.10</td>
<td>47.46</td>
</tr>
<tr>
<td></td>
<td>N = 1034</td>
<td>(0.69)</td>
<td>(0.71)</td>
<td>(0.66)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>Minority</td>
<td>Reduced or free lunches</td>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
<td>----------</td>
<td>-------------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>52.51</td>
<td>52.29</td>
<td></td>
<td>47.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.64)</td>
<td>(0.73)</td>
<td></td>
<td>(0.52)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>51.58</td>
<td>52.09</td>
<td></td>
<td>46.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.66)</td>
<td>(0.68)</td>
<td></td>
<td>(0.49)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>51.85</td>
<td>51.70</td>
<td></td>
<td>48.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.64)</td>
<td></td>
<td>(0.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>52.96</td>
<td>50.55</td>
<td></td>
<td>49.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.72)</td>
<td></td>
<td>(0.52)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>51.58</td>
<td>52.09</td>
<td></td>
<td>46.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.63)</td>
<td></td>
<td>(0.58)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>51.85</td>
<td>51.70</td>
<td></td>
<td>48.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.59)</td>
<td></td>
<td>(0.52)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>52.96</td>
<td>50.55</td>
<td></td>
<td>47.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.67)</td>
<td></td>
<td>(0.61)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>51.58</td>
<td>52.09</td>
<td></td>
<td>47.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.73)</td>
<td>(0.74)</td>
<td></td>
<td>(0.81)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>51.85</td>
<td>51.70</td>
<td></td>
<td>46.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.78)</td>
<td></td>
<td>(0.95)</td>
<td></td>
</tr>
</tbody>
</table>

\( N_c \) = Number of cases for between subjects variables.

\( N_m \) = Number of measurements for each level of Reduced or Free Lunches, which is a discrete, time-varying covariate.

\( N_t \) = 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.
Table 2. Test Score Marginal Means for Each Program Participation Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Test</th>
<th>Basic Battery</th>
<th>Reading</th>
<th>Language</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matched Controls</td>
<td></td>
<td>47.48&lt;sup&gt;a&lt;/sup&gt;</td>
<td>46.70&lt;sup&gt;a&lt;/sup&gt;</td>
<td>48.05&lt;sup&gt;a&lt;/sup&gt;</td>
<td>48.28&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>N&lt;sub&gt;c&lt;/sub&gt; = 1100</td>
<td></td>
<td>(0.69)</td>
<td>(0.70)</td>
<td>(0.65)</td>
<td>(0.73)</td>
</tr>
<tr>
<td>Program Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N&lt;sub&gt;c&lt;/sub&gt; = 1106</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own Controls</td>
<td></td>
<td>44.98&lt;sup&gt;b&lt;/sup&gt;</td>
<td>45.17&lt;sup&gt;a&lt;/sup&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>46.50&lt;sub&gt;ab&lt;/sub&gt;</td>
<td>45.81&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.74)</td>
<td>(0.77)</td>
<td>(0.74)</td>
<td>(0.82)</td>
</tr>
<tr>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 854</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 857</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 782</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 779</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Only</td>
<td></td>
<td>49.52&lt;sup&gt;c&lt;/sup&gt;</td>
<td>49.37&lt;sup&gt;c&lt;/sup&gt;</td>
<td>49.77&lt;sub&gt;ac&lt;/sub&gt;</td>
<td>50.30&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.71)</td>
<td>(0.74)</td>
<td>(0.70)</td>
<td>(0.78)</td>
</tr>
<tr>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 1179</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 1179</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 1181</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 1170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous Only</td>
<td></td>
<td>42.54&lt;sup&gt;d&lt;/sup&gt;</td>
<td>42.34&lt;sub&gt;d&lt;/sub&gt;</td>
<td>45.67&lt;sub&gt;abd&lt;/sub&gt;</td>
<td>38.74&lt;sub&gt;d&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.99)</td>
<td>(1.06)</td>
<td>(1.09)</td>
<td>(1.16)</td>
</tr>
<tr>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 197</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 197</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 197</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 195</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current &amp; Previous Year</td>
<td></td>
<td>54.50&lt;sup&gt;e&lt;/sup&gt;</td>
<td>53.59&lt;sup&gt;e&lt;/sup&gt;</td>
<td>53.98&lt;sup&gt;e&lt;/sup&gt;</td>
<td>53.92&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.12)</td>
<td>(1.21)</td>
<td>(1.23)</td>
<td>(1.31)</td>
</tr>
<tr>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 203</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 205</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 205</td>
<td>N&lt;sub&gt;m&lt;/sub&gt; = 203</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Marginal means within a column that do not share common subscripts differ significantly at p < .05 using Holm’s (Holm, 1979) sequential Bonferroni test. Standard errors of means are shown in parentheses.

N<sub>c</sub> = Number of cases for between subjects variables.

N<sub>m</sub> = Number of measurements for each Program Participation condition.
Footnotes

1. Estimated marginal means for the levels of each demographic grouping variable were computed by creating an appropriate linear combination of the model parameter estimates involved in the effect, holding all other parameters constant at their “neutral values.” For example, marginal means for females were computed as 1.0 times the intercept parameter estimate, plus 1.0 times the Gender parameter estimate (since Females were coded as 1’s on the Gender variable), plus .5 times the Ethnicity parameter estimate, plus .5 times the Free or Reduced-Price Lunches parameter estimate, plus 10.0 times the Age parameter estimate (since the average age of students in elementary schools is usually close to ten years). The value of .5 was selected as the neutral value for each 0/1-coded demographic variable so that the computed marginal means would be unweighted estimated population marginal means, and would not reflect the specific distributions of group membership on these variables found in the present sample. (For a thorough discussion of estimated population marginal means, see Searle, Speed., & Milliken, 1980).
References


Violence prevention


