Fonagy, P; Twemlow, SW; Vernberg, EM; Nelson, JM; Dill, EJ; Little, TD; Sargent, JA; (2009) A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. **J CHILD PSYCHOL PSYC** , 50 (5) 607 - 616. 10.1111/j.1469-7610.2008.02025.x.

RESEARCH ARTICLE

A Cluster Randomized Controlled Trial of Child-Focused Psychiatric Consultation and a School Systems-Focused Intervention to Reduce Aggression

Peter Fonagy[12], Ph.D.; stuart W. Twemlow[2], M.D.; Eric M. Vernberg[3], Ph.D.; Jennifer Mize Nelson[3], M.A.

Edward J. Dill[3], Ph.D.

Todd D. Little[4], Ph.D.

John A. Sargent[2], M.D.

[1] University College London
[2] Menninger Department of Psychiatry, Baylor College of Medicine
[3] Clinical Child Psychology Program, University of Kansas
[4] Quantitative Psychology Program, University of Kansas

Peter Fonagy, Sub-Department of Clinical Health Psychology, University College London, Gower Street, London WC1E 6BT, p.fonagy@ucl.ac.uk

ABSTRACT

Background: While school-based anti-bullying programs are widely used, there have been few controlled trials of effectiveness. This study compared the effect of manualized School Psychiatric Consultation (SPC), CAPSLE (a systems and mentalization focused whole school intervention), and treatment-as-usual (TAU), in reducing aggression and victimization among elementary school children.

Method: 1,345 third to fifth graders in nine elementary schools in a medium-sized Midwestern city participated in a cluster-level randomized controlled trial with stratified restricted allocation, to assess efficacy after two years of active intervention and effectiveness after one year of minimal input maintenance intervention. Outcome measures included peer and self reports of bullying, bystanding, and mentalizing behavior and classroom behavioral observations of disruptive and off-task behavior.

Results: CAPSLE moderated the developmental trend of increasing peer-reported victimization (p<.01), aggression (p<.05), self-reported aggression (p<.05) and aggressive bystanding (p<.05), compared to TAU schools. CAPSLE also moderated a decline in empathy and an increase in the percent of children victimized compared to SPC (p<.01) and TAU conditions (p<.01). Results for self-reported victimization, helpful bystanding, and beliefs in the legitimacy of aggression did not suggest significantly different changes among the study conditions over time. CAPSLE produced a significant decrease in off-task (p<.001) and disruptive classroom behaviors (p<.01), while behavioral change was not observed in SPC and TAU schools. Superiority with respect to TAU for victimization (p<.05), aggression (p<.01), and helpful (p<.05) and aggressive bystanding (p<.01) were maintained in the follow-up year.

Conclusions: A teacher-implemented school-wide intervention that does not focus on disturbed children substantially reduced aggression and improve classroom behavior.

Keywords: childhood aggression, anti-bullying intervention, psychiatric consultation, mentalization, randomized controlled trial

Abbreviations: SPC (School Psychiatric Consultation); CAPSLE (Creating a Peaceful School Learning Environment)

Bullying has broad effects on children's mental health (Smith & Sharp, 1994), including early disruptive and aggressive behavior (Nansel et al., 2003), school dropout, substance abuse (Kumplainen and Rasanen, 2000), depressed mood, anxiety, and social withdrawal (Dill et al., 2004; Shafii and Shafii, 2003; Swearer et al., 2004). It also undermines educational achievement (Greenberg et al., 2003) and disrupts children's abilities to develop social relationships (Masten and Coatsworth, 1998).

Meta-analyses of over 300 school-based violence intervention programs (Mytton et al., 2002; Wilson et al., 2001; Wilson et al., 2003) suggest that effective programs are research-based (ES=.24-.36), small high-risk sample, single group designs with highly-trained teachers and reactive measures such as therapy counseling, behavioral classroom management, and social competence enhancement. Programs directly targeting aggressive behavior are no more effective than those focusing on other aspects of social relationships. Though there have been more individual- than environmental-focused interventions (ratio of 4:1), both appear equally effective (Smith, Ananiadou, & Cowie, 2003).

Our study contrasts two school-wide interventions with a treatment-as-usual control group in a cluster-randomized longitudinal trial with one-year post-intervention follow-up. School Psychiatric Consultation (SPC) is a manualized protocol aimed at addressing mental health issues of children with disruptive behavioral problems, internalizing problems or poor academic performance. Thirty-five consultation outcome studies utilizing programs similar to SPC showed improvement in academic performance (Berkovitz, 2001) and positive changes for

children (Pearson et al., 2001). SPC represents individually-tailored interventions targeting children with notable adjustment problems.[1]

CAPSLE (Creating a Peaceful School Learning Environment), also manualized, is a psychodynamic social systems approach addressing the co-created relationship between bully, victim, and bystanders (Twemlow et al., 2004; Twemlow et al., 1996) in the tradition established by Dan Olweus (1996) and Peter K. Smith (Smith & Sharp, 1994). It assumes that all members of the school community, including teachers, play a role in bullying. It aims to improve the capacity of all community members to mentalize, that is, to interpret both one's own and others' behaviour in terms of mental states (beliefs, wishes, feelings), assuming that greater awareness of other people's feelings will counteract the temptation to bully others. In contrast to SPC, CAPSLE represents a whole-school intervention approach. Details of the strategies used to enhance mentalization are given in the supporting materials.

In a pilot investigation in a high-risk elementary school CAPSLE reduced the number of disciplinary referrals for aggressiveness and improved achievement test scores (Twemlow et al., 2001).

We expected to find lower levels of aggression and victimization over time in schools receiving SPC and CAPSLE, compared to schools in the TAU condition. However, because CAPSLE, but not SPC, systematically addresses power dynamics and mentalizing about bullying, we predicted more positive bystanding behaviors, greater empathy for victims, and less favorable attitudes towards aggression only in CAPSLE schools. Finally, we predicted that effects for CAPSLE during the active implementation phase (Years 1 and 2) would be sufficiently incorporated into the school ethos to be sustained in the absence of active training

and supervision and thus would be maintained in Year 3 with minimal research input, and that these effects on aggression and victimization would be similar to those produced by ongoing SPC.

Method

This was a cluster randomization trial (ISRCTN 15430198), where schools served as clusters and were randomly assigned to intervention conditions, but the unit of inference remained at the individual level (Campbell et al., 2001; see Figure 1 in main text and Figure 2 in supplemental materials). Elementary schools (Kindergarten-5th grade) were recruited by presenting the research plan to principals. All 19 schools in the district were eligible and 10 agreed to participate. Prior to random assignment to condition, a stratified restricted allocation procedure was used to stratify schools based on percentage of low-income students indicated by students' free- and reduced-lunch status. Randomization was by a statistician independent of the study. Resource constraints allowed no more than three schools in the SPC and CAPSLE conditions, leaving four schools for the TAU condition. The principal of one school assigned to TAU withdrew upon learning of this assignment. Following randomization, children did not differ across study conditions on baseline demographic characteristics (see Table 1 in supportive documentation). In the experimental conditions (CAPSLE and SPC) full intervention was offered for two years (the efficacy phase) with a limited third year of intervention designed to mimic implementation outside of a research protocol by reducing the intensity of both CAPSLE and SPC (the effectiveness phase). Baseline outcome measures were gathered in the Fall of the first school year immediately before the start of intervention. Intervention efficacy was tracked from baseline through Time 4 (Years 1

and 2); the effectiveness phase was then assessed based on data from Times 5 and 6 (Year 3).

All 3rd-5th graders in the schools were eligible to participate. Inclusion in the outcomes research component of this trial required signed parental permission and child assent each year, obtained by distributing at enrollment a detailed description of the study and permission form to parents. Teachers prompted return of forms and offered incentives for classes that achieved a 70% return rate, regardless of whether return forms allowed participation. Over three school years, 74% of eligible children received permission to participate. Table 5 in the supplementary materials includes information on the proportion of children in each school in each of the three years who participated in providing the outcome measures. Participants were predominantly low-income (64%) and lived in single-parent households (57%). Minorities were well-represented (42%). Comparisons of demographic information for participants and non-participants indicated no consistent differences across years or treatment conditions, with the exception of lower overall participation of low-income children in the treatment-as-usual condition only.

Interventions

School Psychiatric Consultation. This was a school-level intervention focused on individual children within the schools randomized to this arm of the trial. Three child psychiatry residents, supervised biweekly by a senior child psychiatrist, delivered mental health consultation following the SPC manual for four hours a week throughout the first two school years. The psychiatry residents attended weekly school resource meetings and consulted directly with teachers, parents and other school personnel, through classroom observations and meetings, providing 140 consultations for 65 students in Year 1 and 97 consultations for 45 students in Year

2. Interventions typically included modifications in behavior management strategies for school staff and parents and recommendations for individual or group counseling. A more detailed breakdown is included in the supplemental materials. In the third, post-demonstration year of the study, in order to prevent group differences arising merely from service withdrawal, two child psychiatrists continued to provide unsupervised consultation following the SPC manual. The same professionals were employed for all schools in the SPC arm.

CAPSLE. This was a school-wide intervention that aimed to modify educational and disciplinary school climate. A CAPSLE team drawn from school staff in the pilot project led implementation in Years 1 and 2 using a training manual (Twemlow et al., 1999). In Year 1, teachers received a day of group training, students received 9-sessions of of self-defence training, and the CAPSLE team consulted with school staff monthly. Year 2 began with a school-wide half-day refresher training for all school staffand a 3-session refresher self-defence course, and consultation continued to counselors, teachers, and the adult/peer mentor programs. The second author led biweekly supervision meetings with the intervention team during Years 1 and 2. At the beginning of Year 3, the CAPSLE team ceased working with the schools, and only in-service refresher training was provided to school staff. Self-defence training continued as in Year 2. Intervention fidelity was assessed using a teacher self-report measure that required teachers to state the frequency with which various CAPSLE program components were implemented and, through an assessment of teaching attitudes, identified the extent of a teacher's consistency with CAPSLE principles. While adherence level varied, over 80% of teachers claimed to implement at least 3 out of the 5 CAPSLE

components. In a separate paper, we report relations between teacher adherence to CAPSLE and child outcomes (Biggs, Vernberg, Twemlow, Fonagy, & Dill, in press).

Data Collection

Reports of aggression, victimization, bystanding behavior, and mentalizing were gathered twice yearly (October-November, March-April). Trained research assistants administered questionnaires to classrooms of children in three 15- to 45-minute sessions one week apart during each data collection period. Research assistants conducted behavioral observations at mid-year (January, February) on a randomly chosen subgroup of participating children each year of active intervention (three boys and three girls randomly selected from a 3rd, 4th, and 5th grade classroom in each of the nine participating schools, total *n*=162 each year).

Measures

Peer nominations of aggression, victimization, and bystanding. Children were asked to circle the names of classmates fitting each of 6 items describing overt and relational aggression and 6 items describing overt and relational victimization (Crick and Bigbee, 1998), as well as 3 items each describing aggressive and helpful bystanding (Vernberg et al., 1999). Children could nominate as many classmates as desired for each of the 18 items (Perry, 1988). A child's score for each construct was the proportion of classmates who nominated them. The average for each construct (*Total Aggression-Peer, Total Victimization-Peer, Aggressive Bystanding, Helpful Bystanding)* was calculated with all Cronbach's alphas >.80 across all six measurement points.

Self reports of aggression, victimization, and mentalizing. The Peer Experiences Questionnaire (Vernberg et al., 1999) was used to obtain self-reports of children's experiences as targets and perpetrators of aggression, and concern for

victims' experience and beliefs about aggression on 5-point Likert scales. All four scales demonstrated adequate internal consistency at all six timepoints (alpha>.65).

Observations of classroom behavior. Each child was observed for twenty 30-second intervals on three different days (10-minute samples each day, for a total of 60 intervals) by observers blind to the study hypotheses and trained to adequate levels of reliability using classroom observation procedures (Lochman, 1992; Milich and Fitzgerald, 1985). Observers coded off-task behavior, disruptive behavior, and teacher redirections as present or absent for each 30-second interval. Inter-rater reliability was good for each of the three behaviors (kappa = .73, .81, and .93, respectively; single measure intraclass correlations were .87, .87, and .94). The intraclass correlations for the observation variables across the three days ranged from .67 to .69, suggesting they provide reliable indicators of these aspects of the child's experience in the classroom.

Data Analysis

Treatment of missing data. Table 2 in the supplemental materials presents a comparison of the demographic characteristics by trial condition between children with complete data and those who had some level of missing data. Multiple EM imputation (Rubin, 1996) was implemented using SAS PROC MI to estimate the missing data points utilizing the full item pool for participants with at least 2 sets of observations and described in the supportive documentation. Missing data estimation involved 25.3% of the sample. To capture the random variability around "true" values, the set of missing data points was estimated five times, thus creating five datasets. Results presented here were averaged across all datasets using SAS PROC MIANALYZE, which utilizes Rubin's rules for combining results across multiple imputations (www.SAS.com).

Statistical procedures. SAS PROC MIXED was used to conduct hierarchical linear modeling (HLM) analyses because of the measurement of within-subject change on the study outcomes across multiple timepoints (Bryk & Raudenbush, 1987; Singer, 1998). These analyses assessed differential effects of CAPSLE, SPC, and TAU on the outcome variables over time, while controlling for the effects of student gender and family income. Outcomes were assessed in two different phases of HLM analyses: (1) longitudinally starting with the baseline and across the 3 timepoints of measurement during active intervention in Years 1 and 2, and (2) longitudinally across the 2 consecutive semesters of less intense intervention during the 3rd follow-up year. HLM calculates the best-fitting slope through all timepoints of data utilized (i.e., growth in each outcome over the multiple time-points from baseline through active intervention rather than a simple measure of change from baseline to T4). Assessment of change in outcomes during the follow-up year was similarly conducted by calculating a best-fitting line between T5 and T6, to assess whether the levels of the outcome variables at the start of the third year were maintained or changed throughout the year of less intense intervention. The flow of participants in and out of the study was assessed but the number of patterns precluded deriving stable estimates of parameters. Given that flow patterns were randomly distributed across conditions, any bias related to flow in and out of the study would contribute to the error variance used to estimate the significance of the key parameter estimates testing the intervention. Moreover, the flow in and out between T1 (baseline) and T2 (1st assessment of intervention) was minimal (2.9%). School-level differences were not determined to be significant predictors of outcome variables in preliminary analyses and thus were not included in any of the final models.[2]

In separate analyses of classroom observational data, LISREL was used to conduct hierarchical linear modeling (HLM) analyses because of the inherent nesting present when analyzing individual children in the context of classrooms of varying intervention conditions (Singer, 1998). This enabled testing of whether classrooms in each of the three study conditions varied in the degree to which children's behavior changed from Year 1 to Year 2, while controlling for the structure of the data where children shared variance with classmates in their classrooms.

Results

A series of HLM analyses tested the differential effects of CAPSLE and SPC in comparison to the effects of TAU across the two years of active intervention and separately during the third follow-up year. Results from the active intervention years are presented first, followed by a briefer overview of results during the follow-up year. Estimates of key model parameters and their respective effect sizes[3] and significance levels for the intervention years are presented in the top panel of Table 3 (complete set of parameters are available with supplemental materials in table 3a and 3b), with mean scores in Table 4 for each of the eight outcomes of interest. In addition to the *CAPSLE* and *SPC* main effects, the *time* $\times$ *CAPSLE* and *time* $\times$ *SPC* interaction effects are the most pertinent to understanding the differential influence of intervention conditions.

Sample characteristics and main effects. A primary source of variability for most of the variables was time. Across the three groups, most measures showed an increase over the first two years of the study (peer reported aggression, victimization, and bystanding, all $p<.001$) and empathic mentalizing showed a decline ($p<.01$). There are a variety of explanations for these main effects (e.g. changes in the classroom population, secular trends within the school system), but

as our hypotheses are tested by examining the time by condition interaction terms, they do not pose interpretational problems.

There were main effects indicating lower overall levels of self-reported aggression for children in both CAPSLE and SPC schools (p<.05 for both) and a further main effect of CAPSLE indicating an overall difference in helpful bystanding (p<.01). However, in the absence of interactions with time, these differences are conservatively considered as differences in baseline rather than as intervention effects. Regarding other control variables in the models, there were main effects but no interactions with intervention for low income and gender, suggesting that low SES and male gender were associated with higher aggression and lower levels of helpful bystanding.

Intervention by time interactions. The contrast with TAU yielded only one significant interaction with time for children randomized to SPC schools. Helpful bystanding increased somewhat over the course of the trial, but this was evident only in the first year. By contrast CAPSLE showed significant improvement relative to TAU across time on four out of the eight primary outcome variables: peer-reported aggression (p<.05), peer-reported victimization (p<.01), aggressive bystanding (p<.05) and empathic mentalizing (p<.01).

Changes in the outcome variables across the intervention years did not differ significantly between the CAPSLE and SPC schools, except that children in SPC schools reported a significantly larger decrease in self-reported empathy over time compared to children in CAPSLE schools ($t_{(adj.183.79)}$=-2.90,p<.01,d=.19).

A standard cutoff of one standard deviation above the mean level of peer-reported victimization across trial conditions at baseline and again at Time 4 was used to categorize children as either victimized or non-victimized (see Table 4 in

supplemental materials) At baseline the groups did not differ but at Time 4 the percentage of children victimized was less in the CAPSLE than in the TAU condition (19.2% vs. 26.1%), RR=.73, (95% CI: .58 , .98), $\chi^2(1)=5.7$, p<.02. and less in CAPSLE than in the SPC condition (19.2% vs. 24.9%), RR=.77, (95% CI: .61, .98), $\chi^2(1)=4.3$, p<.04.

Follow-up Analyses. Secondary analyses were conducted to assess the maintenance of intervention effects during the follow-up year (see last panel of Table 3). The main effects for CAPSLE and SPC contrasted with TAU illustrate the degree to which outcomes were maintained in the third year. The CAPSLE main effect (p<.05) indicates that children in CAPSLE schools, but not in SPC schools, experienced significantly less victimization in the fall of the third year compared to TAU schools, even though at baseline students in CAPSLE schools exhibited greater victimization compared to TAU schools. During the follow-up year, children in CAPSLE schools also continued to experience significantly less peer-reported aggression than children in TAU schools (p<.01) and more helpful bystanding (p<.05) while children in SPC schools were not significantly lower (p<.10). Children in both CAPSLE (p<.01) and SPC (p<.05) schools reported significantly less aggressive bystanding compared to TAU schools. Empathy remained relatively stable only in CAPSLE schools, whereas TAU schools evidenced a decline in levels of empathy over the three-year period (p<.05).

The comparisons of SPC and CAPSLE in the follow-up year indicated that in the fall of the third year, children in SPC schools displayed significantly less helpful bystanding ($t_{(adj.446.7)}$=-2.70, p<.01,d=.18), more self-reported victimization ($t_{(adj.134.72)=}$2.79, p<.01,d=.18) and perceived aggression as more legitimate ($t_{(adj.120.88)=}$3.21, p<.01,d=.21) compared to children in CAPSLE schools.

Classroom Behavioral Observations. Three separate series of HLM analyses were conducted to evaluate the effects of the intervention conditions on log-transformed classroom-level proportions of observed off-task and disruptive classroom behaviors and teacher redirections across the two years of active intervention (classroom-level ICCs = 0.33-0.52 across outcome variables). Results indicated significant classroom-level interactions between study year and dummy codes contrasting CAPSLE with SPC and TAU for both off-task (CAPSLE vs. TAU: β = -0.16, z = -5.04, p<.001; SPC vs. CAPSLE: β = 0.14, z = 4.43, p<.001) and disruptive behaviors (CAPSLE vs. TAU: β = -0.06, z = -2.67, p<.01; SPC vs. CAPSLE: β = 0.08, z = 3.47, p<.001). Procedures outlined by Preacher, Curran, and Bauer (2006) were utilized to probe these two-way interactions. In classrooms receiving the CAPSLE intervention children showed a decline in off-task (simple slope = -0.153, 95% CI: -0.198, -0.108) and disruptive behaviors (simple slope = -0.070, 95% CI: -0.103, -0.036) from Year 1 to Year 2. However, children in SPC and TAU classrooms showed little or no differences in off-task (SPC simple slope = -0.014, 95% CI: -0.056, 0.028; TAU simple slope = 0.003, 95% CI: -0.038, 0.043) and disruptive behaviors (SPC simple slope = 0.012, 95% CI: -0.019, 0.044; TAU simple slope = -0.008, 95% CI: -0.038, 0.023) across the active intervention years. With respect to observed teacher redirections, results indicated significant classroom-level interactions between study year and dummy codes contrasting TAU with CAPSLE and SPC (CAPSLE vs. TAU: β = 0.07, z = 3.64, p<.001; SPC vs. TAU: β = 0.04, z = 2.05, p<.05). Probing these interactions, we found that children in TAU classrooms were given significantly less teacher redirections in Year 2 compared to Year 1 (simple slope = -0.042, 95% CI: -0.067, -0.018). Children in CAPSLE schools experienced an increase in teacher redirections in Year 2 (simple slope = 0.028,

95% CI: -0.001, 0.057), and children in SPC schools remained relatively stable regarding the teacher redirections received over time (simple slope = -0.005, 95% CI: -0.031, 0.021). Second year intervention versus TAU group observations yielded the following effect sizes for off-task and disruptive behaviors: 1.1 (95% CI: -.9, 3.1) and .84 (95% CI: -1.2, 2.8) for CAPSLE and .29 (95% CI: -1.7, 2.3) and .63 (95% CI: -1.4, 2.6) for SPC.

Discussion

Our major finding is that a simple mentalization/power dynamics focused anti-violence program (CAPSLE) provides an effective teacher-administered protocol relative to no intervention in reducing children's experience of aggression and victimization. CAPSLE's effectiveness was indicated by reduction in the number of children nominated by their peers as aggressive, victimized, or engaging in aggressive bystanding. This was confirmed by behavioral observation of reduced disruptive and off-task classroom behavior in CAPSLE schools. In the third year (the maintenance phase) CAPSLE remained superior to no intervention in peer-reported aggression and victimization. It is unlikely that the superiority of CAPSLE over no treatment can be attributed to the mere presence of an intervention in these schools. SPC also modified schools' approach to mental health problems and was enthusiastically received by school staff but had limited measurable impact on the general level of aggression and victimization.

As both behavioral observation and peer nominations reflected an effect in CAPSLE schools, it is unlikely that these can be attributed to changes in the criteria for ratings introduced by the program. Possibly children were less objective in describing their own experiences than when reporting on the behavior of others, probably due to social desirability and self-protective biases (Ladd and

Kochenderfer-Ladd, 2002). Notably the only report concerning ‘own behavior’ to yield significant group differences was empathetic mentalizing of victims. This measure may be less obviously prone to a ‘faking good’ bias and hence it registered the impact of the CAPSLE intervention.

The findings on bystanding are puzzling. Aggressive bystanding, as reported by peers, declined and helpful bystanding increased as predicted relative to the non-intervention and SPC group in the CAPSLE condition. This remained steady across the entire period of the project. However, helpful bystanding increased dramatically in the second year in the non-intervention group. This isolated improvement in some control classes may be due to change of teaching staff, leakage of some of the intervention principles to the TAU schools or a combination of these factors.

Table 3 contains estimates of effect sizes based on variances for each timepoint and estimates of mean differences based on the beta coefficients. .These indicate small to moderate effect sizes for the primary outcome variables (mean significant ES=.32, range.20-.59) and medium to strong effects for classroom observations (mean ES=.97, range: .84-1.1). Past research on school-wide multimodal interventions has only demonstrated modest effects on aggression against TAU (Wilson et al., 2003). Medium effect sizes are normally only associated with intense studies with small samples (Wilson et al., 2001).(Smith et al., 2003) CAPSLE produced a number of modest and some large effect sizes, particularly with behavioral measures. The findings are also notable because the sample was not particularly high-risk and high-risk samples normally generate larger effects (Mytton et al., 2002; Wilson et al., 2001; Wilson et al., 2003). The findings are consistent with previous studies that teacher-delivered school-based programs are more effective than those delivered by professionals, researchers, lay adults or peer mentors

(Wilson et al., 2003). We have shown that a program that is not focused on aggressive children, but rather on other aspects of social relationships, has significant effects that may have theoretical as well as practical importance (Farrington, 2003) as multimodal interventions normally yield quite small effects (Wilson et al., 2001).

The relative ineffectiveness of SPC by no means indicates the lack of value of such interventions for individual children. Based on previous findings (Pearson et al., 2001) we expect psychiatric consultation to improve service availability and effectiveness for particular children, but it appears that treating problem children has a relatively low impact on aggression in these schools. CAPSLE enhances school-wide awareness of the omnipresence of power struggles and their effects on the capacity to think about others' points of view. The findings suggested that empathic mentalization was enhanced in CAPSLE schools. They are consistent with the view that the emotional and cognitive skills learned in handling interpersonal power struggles enhance both the emotional and cognitive empathic aspects of mentalizing and self agency (Baron-Cohen, 2005; Blair, 2005) and thus may reduce the likelihood of resorting to physical aggression (Fonagy, 2003). .

In evaluating the results several limitations should be noted. First, the number of schools per condition was relatively small. Although the modeling approach enabled us to independently identify the impact of individual schools, the randomized units may not have been fully equivalent at baseline due to unmeasured factors such as receptivity to the assigned intervention condition. Replication with a larger sample would be important to establish the generalizability of the findings and the possible role of school-level factors in implementation and outcomes. Second, the significant impact of time on children's reports of aggression and victimization remains unclear

but reflects a gradually worsening situation within a school system experiencing resource shortages or a general developmental trend, beginning in third grade, for children to report more aggression toward others, less empathy for victims, and more positive attitudes towards aggression as they progress towards high school.

Third, we included no 'objective' measures of school discipline (e.g., suspensions, expulsions, truancy or other data reported by principals to the school district). Such data are readily available and favor the CAPSLE condition, but since CAPSLE introduced a disciplinary code requiring teachers to reduce disciplinary referrals, we considered these indicators fundamentally contaminated by the experimental manipulation. Fourth, as with any longitudinal school-based study, we had over one quarter of participants with incomplete data which could threaten generalizability. We used multiple imputations with the full study item pool so that effects of variables associated with missingness would be accounted for in the imputed data, and the results are thereby generalizable to the original sample with regard to these variables. The remaining limitation, here, is the possibility that there were other variables associated with missingness that were not in our dataset and the bias from these cannot be corrected. Fifth, we do not know if the intervention works at all levels of aggression or only with the low to moderate levels of victimization (12-17% reporting being regularly victimized) characteristic of this sample. Sixth, the program's overall impact was limited with the majority of effect sizes low, some of the measures (especially those based on first person reporting) revealing no treatment effects and the positive results on the observational measures needing to be qualified by the fact that observers were not blind to intervention condition, owing to the public nature of the positive school climate campaign used in CAPSLE. Finally, the HLM data analytic approach taken involves

the estimation of marginal means, which accounts for the influence of a range of control variables but necessarily obscures the participants' observed scores.

Conclusion

This is one of the first randomized controlled trials to show that an easily implemented school-wide, systems-focused intervention that emphasizes mentalization and power dynamics can reduce children's experiences of aggression in school and improve observed classroom behavior. Although the measured effects on aggression are modest, the effects of CAPSLE relative to no intervention are robust and relative to an active intervention (SPC), small but significant. The impact of a school-wide intervention may occur at multiple levels, such as improving school morale and improving the classroom learning environment by decreasing tensions and negative emotions that accompany bully-victim problems. Prior research has shown a clear improvement in academic performance for children who spent two or more years in schools offering this program (Fonagy et al, 2005). Combining this whole-school bully-victim-bystander program with more traditional child-focused consultation may amplify the impact of both forms of intervention.

Figure Caption

Figure 1 [see Powerpoint file]. All children, grades K-5, participated in the interventions but only those in grades 3-5 completed measures because of the developmental level required. The lower portion of the diagram indicates the number of 3rd-5th grade children who completed study outcome measures in Years 1 and 2 (full intervention years, Time 1 of Year 1 = baseline) and Year 3 (the effectiveness phase). The majority of the sample was comprised of the same children from year to year (i.e., 3rd graders in Year 1 were 5th graders in Year 3); however, increases and decreases in the number of children within each intervention condition across the three years was due to mainly difference in size between the 5th grade cohort that aged out of the study at the end of one year and the new 3rd grade cohort in the following year, but also to changes in the number of children with consent to complete the outcome measures each year, and residential relocation that affected children moving into and out of the study schools (for full account see supplementary materials). Multiple imputation capitalized on the information gained from children when they did participate and used this information to maximize the total number of cases used in analyses assessing change across the active intervention years and the follow-up year (see bottom of diagram for total numbers of cases included in analyses, which represent the number of children in each condition for whom data was collected in at least 2 of the study semesters).

Fonagy, P; Twemlow, SW; Vernberg, EM; Nelson, JM; Dill, EJ; Little, TD; Sargent, JA; (2009) A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. **J CHILD PSYCHOL PSYC** , 50 (5) 607 - 616. 10.1111/j.1469-7610.2008.02025.x.

Table 3: Parameter Estimates (and ES's) for HLM Analyses across Two Years of Active Intervention and One Year of Follow-Up (fuller version in supportive materials)

Parameter	Peer-Report Aggression	Self-Report Aggression	Peer-Report Victimization	Self-Report Victimization	Aggressive Bystanding	Helpful Bystanding	Empathy	Aggression is Legitimate
CAPSLE	.09(.01)	1.59(.15)*	1.17 (.12)†	.30(.12)	.57(.05)	2.22(.20)**	-1.93(.17)**	2.47(.21)**
SPC	-.54(.05)	1.73(.15)*	.21(.02)	.99(.10)	-.39(.04)	-.18(.02)	-.71(.06)	3.22(.28)***
Time × CAPSLE	-.89(.25)*	-.37(.30)*	-.97(.30)**	.22(.06)	-.73(.20)*	-2.37(.59)***	1.01(.26)**	-.57(.09)
Time × SPC	-.43(.11)	-.23(.11)	-.48(.11)	.14(.04)	-.21(.05)	-2.15(.57)***	.04(.01)	-.08(.06)
				One Year Follow-Up				
CAPSLE	-1.91(.20)**	-.47(.04)	-1.99(.20)*	-.79(.08)	-2.24(.21)**	1.94(.17)*	-.63(.06)	-.56(.06)
SPC	-1.63(.13)†	-.13(.01)	-.92(.09)	1.33(.10)	-2.11(.17)**	.11(.01)	.48(.,04)	1.74(.13)†
Time × CAPSLE	-.24(.02)	-.29(.01)	-1.05(.07)	.02(.00)	.07(.00)	-2.53(.15)*	1.21(.08)	-.46(.03)
Time ×	-1.41(.13)†	-.22(.01)	-1.01(.07)	-.85(.04)	.80(.05)	-3.03(.18)**	.21(.01)	-.44(.02)

SPC

Note: Scores were converted to t scores (*M*=100, *SD*=10). ES=effect size (see footnote 3 for method of calculation)

*p<.05. **p<.01. ***p<.001. † indicates marginal significance at p<.10.

Table 4: Means for Outcome Variables in Each Intervention Condition over Time

	Active Intervention Years				Follow Up Year	
Intervention	Baseline	Time 2	Time 3	Time 4	Time 5	Time 6
			Aggression			
Peer Report						
TAU	97.8(.48)	99.6(.61)	101.0(.59)	102.7(.72)	99.5(.54)	103.2(.63)
CAPSLE	98.2(.43)	99.9(.52)	99.8(.46)	101.7(.47)	97.7(.41)	101.2(.41)
SPC	97.5(.44)	99.6(.52)	100.2(.55)	101.6(.65)	98.4(.48)	100.7(.55)
Self Report						
TAU	98.2(.56)	99.3(.57)	99.0(.55)	99.7(.59)	99.3(.49)	100.9(.66)
CAPSLE	100.4(.49)	99.7(.50)	100.0(.47)	100.2(.49)	98.9(.42)	100.3(.46)
SPC	100.6(.53)	100.2(.52)	100.3(.59)	101.1(.61)	99.5(.50)	101.0(.63)
			Victimization			
Peer Report						
TAU	97.6(.56)	99.1(.73)	100.2(.67)	102.8(.74)	100.0(.68)	102.8(.75)
CAPSLE	98.7(.41)	99.9(.44)	100.1(.39)	100.7(.39)	98.0(.44)	99.8(.42)
SPC	97.8(.55)	100.5(.54)	100.3(.56)	101.9(.63)	99.3(.46)	101.1(.45)
Self Report						
TAU	99.70 (.61)	101.0(.71)	98.7(.53)	99.9(.61)	99.8(.56)	100.2(.59)
CAPSLE	100.64	99.4(.46)	99.1(.46)	99.2(.44)	99.0(.45)	99.4(.43)

	(.46)					
SPC	100.63	101.0(.58)	100.2(.58)	100.6(.61)	101.1(.55)	100.7(.60)
	(.54)					

Table 4: Means for Outcome Variables in Each Intervention Condition over Time (continued)

	Active Intervention Years				Follow Up Year	
Intervention	Baseline	Time 2	Time 3	Time 4	Time 5	Time 6
			Aggressive Bystanding			
TAU	97.6(.50)	99.4(.66)	100.3(.58)	102.7(.71)	100.2(.56)	102.2(.65)
CAPSLE	98.1(.44)	100.1(.55)	100.4(.46)	101.2(.45)	98.1(.41)	100.2(.50)
SPC	97.1(.41)	100.0(.49)	100.5(.52)	102.2(.63)	98.6(.47)	101.5(.57)
			Helpful Bystanding			
TAU	96.6(.54)	100.2(.64)	104.0(.69)	104.0(.60)	98.5(.58)	102.3(.64)
CAPSLE	99.4(.48)	100.7(.51)	100.5(.43)	100.3(.48)	100.2(.48)	101.4(.50)
SPC	96.7(.53)	99.6(.54)	98.8(.51)	98.9(.50)	98.2(.45)	98.9(.50)
			Mentalizing			
Empathy						
TAU	102.2(.61)	101.3(.63)	100.4(.59)	98.8(.59)	101.1(.60)	98.8(.59)
CAPSLE	100.4(.47)	99.1(.50)	100.0(.47)	99.1(.47)	100.3(.47)	99.2(.48)
SPC	101.5(.53)	100.3(.55)	99.0(.55)	98.3(.55)	101.3(.51)	99.2(.53)
Aggression is Legitimate						
TAU	96.5(.47)	98.1(.56)	98.7(.52)	99.0(.50)	98.9(.55)	100.4(.59)
CAPSLE	99.1(.50)	100.6(.56)	100.7(.47)	100.9(.49)	98.4(.42)	99.5(.43)

SPC	100.5(.54)	99.9(.53)	101.8(.59)	102.1(.60)	100.9(.54)	102.0(.64)

Note. Numbers in parentheses are standard errors.

Acknowledgements

This research was supported through the Child and Family Center of Menninger Clinic. The authors gratefully acknowledge participating staff and students of USD 501, Topeka, Kansas.

Footnotes

[1]Program manuals for SPC and CAPSLE are accessible at http://www.backoffbully.com.

[2]Three-level HLM models were run as part of preliminary analyses to examine the nesting of children's longitudinal data within schools. The ICCs for this third level averaged 0.04 across all outcome variables of interest. Given the small amount of variance attributable to the school level and the model convergence problems experienced when trying to run these three-level models with predictor variables, school-level variance was not included in the final models.

[3] Effect sizes for main effects were computed by dividing the beta estimate for main effect by the square root of the variance estimate at every timepoint and averaging these to obtain the overall effect size for the group. We estimated the effect size of Group x Time interactions by multiplying the beta estimate for the interaction term with the number of timepoints and dividing by sqrt(var(y_1)+var(y_n) – 2*cov(y_1,y_n)) where var(y_n) is the estimated variance at the 1st and 4th timepoint and cov(y_1,y_n) is the estimated covariance between the two timepoints.

Text for text box

- Overt and relational aggression is common in schools. A range of relatively complex prevention programmes have been shown to be reasonably effective.
- This study evaluates a school-wide readily implemented prevention protocol based on the suggestion that enhancing awareness of one's own and others' thoughts and feelings within the school environment serves to reduce the level of bullying.
- Enhancing mentalization in schools appears to reduce victimization and disruptive behaviour observed in the classroom and increase empathy
- Creating situations where awareness of the subjective experience of other children becomes a focus may be an important component of violence prevention

References

Baron-Cohen, S. (2005). Autism. In *Cambridge encyclopedia of child development* (Vol. 28, pp. 109-126). Cambridge: Cambrdige University Press.

Berkovitz, I.(2001). Evaluations of outcome in mental health consultation in schools. *Child and Adolescent Psychiatric Clinics of North America* 10: 93-103

Biggs, B. K., Vernberg, E. M., Twemlow, S. W., Fonagy, P., & Dill, E. J. (in press). Teacher adherence and its relation to teacher attitudes and student outcomes in an elementary school-based violence prevention program. School Psychology Review.

Blair, R. J. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition, 14*(4), 698-718.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101,* 147-158.

Campbell M, Donner A, Elbourne D (2001), Design analysis of cluster randomized trials. *Statistical Medicine (Special Issue)* 20: 329-496

Crick NR, Bigbee MA (1998), Relational and overt forms of peer victimization: a multi-informant approach. *Journal of Consulting and Clinical Psychology* 66: 337-347

Dill E, Vernberg E, Fonagy P, Twemlow S, Gamm B (2004), Negative affect in victimized children: the roles of social withdrawal, peer rejection, and attitudes to bullying. *Journal of Abnormal Child Psychology* 32: 139-173

Farrington DP, & Coy, J. W. (Eds.) (2003), *Early Prevention of Adult Antisocial Behaviour* Cambridge: Cambridge University Press

Fonagy P (2003), Towards a developmental understanding of violence. *Br J Psychiatry* 183: 190-2

Fonagy P, Twemlow, S. W., Vernberg, E. M., Sacco, F. C., & Little, T. D. (2005), Creating a peaceful school learning environment: The impact of an antibullying program on educational attainment in elementary schools. *Medical Science Monitor* 11: CR317-325

Graham, J. W., Cumsille, P. E., & Elek-Fisk, E. (2003). Methods for handling missing data. In W. F. Velicer & J. A. Schinka (Eds.), *Handbook of psychology: Research methods in psychology, Vol. 2* (pp.87-114). New York, NY:Wiley.

Greenberg M, Weissberg R, O'Brien M, Zins J, Fredewricks L, Resnik H, Elias M (2003), Enhancing school-based prevention and youth development through coordinated social, emotional and academic learning. *American Psychologist* 58: 466-474

Hofer, S. M., & Hoffman, L. (2007). Statistical analysis with incomplete data: A developmental perspective. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 13-32). Mahwah, NJ: LEA.

Kumplainen K, Rasanen E (2000), Children involved in bullying at elementary school age: their psychiatric symptoms and deviance in adolescence, an epidemiological sample. *Child Abuse Neglect* 24: 1567-1577

Ladd GW, Kochenderfer-Ladd B (2002), Identifying victims of peer aggression from early to middle childhood: analysis of cross-informant data for concordance,

estimation of relational adjustment, prevalence of victimization, and characteristics of identified victims. *Psychological Assessment* 14: 74-96

Lochman JE (1992), Cognitive-behavioral intervention with aggressive boys: three-year follow-up and preventive effects. *Journal of Consulting and Clinical Psychology* 60: 426-432

Masten A, Coatsworth J (1998), The development of competence in favorable and unfavorable enviroments. Lessons from research on successful children. *Am. Psychol* 53: 205-220

Milich R, Fitzgerald G (1985), Validation of inattention/overactivity and aggression ratings with classroom observations. *Journal of Consulting and Clinical Psychology* 53: 139-140

Mytton J, DiGuiseppi C, Gough D, Taylor R, Logan S (2002), School based violence prevention programs: systematic review of secondary prevention trials. *Archives of Pediatrics & Adolescent Medicine* 156: 752-762

Nansel T, Overpeck M, Hanie D, Ruan W, Scheidt P (2003), Relationship between bullying and violence in US youth. *Arch. Pediatr. Adolesc. Med.* 157: 348-353

Olweus, D. (1996). Bullying at school: Knowledge base and an effective intervention programme. *Annals of the New York Academy of Sciences, 794*, 265-276.

Pearson G, Jennings J, Norcross J (2001), A program of comprehensive school based mental health services in a large urban public school district: the Dallas model. *Child and Adolescent Psychiatric Clinics of North America* 10: 207-231

Perry DG, Kusel, S. J., & Perry, L. C. (1988), Victims of peer aggression. *Developmental Psychology* 24: 807-814

Preacher KJ, Curran, PJ, & Bauer, DJ. (2006), Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics* 31: 437-448

Rubin DB (1996), Multiple imputation after 18+ years. *Journal of the American Statistical Association* 91: 473-489

Shafii M, Shafii S (2003), School violence, depression and suicide. *Journal of Applied Psychoanalytic Studies* 5: 155-169

Singer JD (1998), Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics* 24: 323-355

Smith, P. K., Ananiadou, K., & Cowie, H. (2003). Interventions to reduce school bullying. *Can J Psychiatry, 48*(9), 591-599.

Smith, P. K., & Sharp, S. (Eds.). (1994). *School bullying: Insights and perspectives*. London: Routledge.

Swearer S, Grills A, Haye K, Cary P (2004), Internalizing problems in students involved in bullying and victimization: Implications for intervention. In: *Bullying in American Schools: A Social-Ecological Perspective on Prevention and Intervention*, Espelage D, Swearer S, eds. New Jersey: Lawrence Erlbaum and Associates

Twemlow S, Fonagy P, Gies M, Evans R, Ewbank R (2001), Creating a Peaceful School Learning Environment: a controlled study of an elementary school intervention to reduce violence. *American Journal of Psychiatry* 158: 808-810

Twemlow S, Sacco F, Twemlow SW (1999), *Creating a Peaceful School Learning Environment: A Training Program for Elementary Schools* Agawam, MA: T & S Publishing Group

Twemlow S, Sacco F, Williams P (1996), A clinical and interactionist perspective on the Bully/Victim/Bystander relationship. *Bulletin of the Menninger Clinic* 60: 296-313

Vernberg EM, Jacobs AK, Hershberger SL (1999), Peer victimization and attitudes about violence during early adolescence. *Journal of Clinical Child Psychology* 28: 386-395

Wilson D, Gottfredson D, Najaka S (2001), School based prevention of problem behaviors: a meta analysis. *Journal of Quantitative Criminology* 17: 247-272

Wilson S, Lipsey M, Devzon J (2003), The Effects of School Based Intervention Programs on Aggressive Behavior: A Meta analysis. *Journal of Consulting & Clinical Psychology* 71: 136-149