Preventing Childhood Aggression and Bullying: Implications for School-Based Intervention Research from the Denver Public Schools Trial

Jeffrey M. Jenson, Ph.D.
University of Denver
Denver, Colorado

William A. Dieterich, Ph.D.
Northpointe Institute for Public Management
Golden, Colorado

October 5, 2009

This manuscript was presented at the Los Angeles Conference on Intervention Research in Social Work. School of Social Work, University of Southern California. Los Angeles, CA. October 23, 2009.
Abstract

Outcomes from a group-randomized trial of a prevention curriculum aimed at preventing bullying and victimization among elementary students in the Denver, Colorado public school system are presented. Twenty-eight elementary schools were randomly assigned to receive selected modules of *Youth Matters*, a skills-training curriculum that targets bullying and victimization, or to a no-treatment control group. Linear growth models were fitted to four waves of data collected over two years to test the effect of the intervention on the rate of change in self-reported bullying and victimization. Participation in *Youth Matters* was associated with a 20 percent decline in bully victimization at the end of the fourth-grade. Practice and methodological challenges encountered in the investigation are discussed in the larger context of intervention research in school settings. Strategies to increase school-based intervention research by social work investigators are outlined.

Keywords: children, adolescents, bullying, victimization, prevention
School-based prevention has been heralded as an efficacious and cost-effective way to decrease the prevalence of childhood and adolescent antisocial behavior (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004; Gottfredson & Wilson, 2003; Hawkins, 2006). Bolstered by evidence indicating that empirically-based curricula can prevent or delay the onset of problems like drug use or delinquency, many schools have strengthened their commitment to providing prevention content at all educational levels (Jenson, in press). Smoking, illicit drug use, delinquency, and health concerns have been common targets of school-based prevention efforts.

In recent years many school districts have also placed financial resources into developing and evaluating prevention programs aimed at reducing bullying and bully victimization (Eisenberg & Aalsma, 2005; Jenson, Dieterich, Rinner, Washington, & Burgoyne, 2006; Rigby, Smith, & Pepler, 2004; Spivak & Prothrow-Stith, 2001). The rather sudden interest in bullying prevention was fostered by a series of school shootings that shook the foundation of American public education in the past decade. Close examination of the incidents of school violence revealed that many perpetrators had been victims of bullying or exclusionary peer practices during childhood and early adolescence (Vossekuil, Fein, Reddy, Borum, & Modzeleski, 2002). These developments, along with the positive outcomes reported in studies of anti-bullying programs in Norway (Olweus, 1993, 1994) and other European countries (Smith, Pepler, & Rigby, 2004), has created considerable interest in bullying prevention programs in many American public school systems.

Prevalence and Nature of Bullying

The word “bullying” is derived from the English language and generally refers to a social process in which a child or an adolescent exerts power or influence over another student in a negative manner to achieve a desired effect or outcome. Olweus (1993) and other investigators
(Frey, Hirschstein, Snell, Edstrom, MacKenzie, & Broderick, 2005; Smith & Brain, 2000) note that bullying occurs in physical, social, and emotional contexts and varies considerably in severity. Overt forms of bullying generally take the form of actual or intended physical harm. Conversely, relational patterns of bullying are most often characterized by social practices that exclude students from desirable peer groups or activities. The concept of a power imbalance in which the perpetrator exerts influence over a victim is an important component and emotional dynamic of bullying (Olweus, 1993; Smith, Morita, Junger-Tas, Olweus, Catalano, & Slee, 1999).

Prevalence estimates of bullying reveal that approximately 30 percent of students engage in overt or relational bullying or are victims of bullying during elementary school (Juvonen, Graham, & Schuster, 2003; Nansel, Overpeck, Pilla, Ruan, Simons-Morton, & Scheidt, 2001; Pellegrini, Bartini, & Brooks, 2001). Students are most likely to be bullied in the classroom or on the playground (Fekkes, Pijpers, & Verloove-Vanhorick, 2005). Boys commit overt or physical acts of bullying at higher rates than girls (Baldry & Farrington, 2000; Nansel et al., 2001). Conversely, girls are more likely than boys to be involved in relational forms of bullying such as excluding others from social activities or spreading rumors (Nansel et al., 2001; Olweus, 1994).

Reports of racial or ethnic differences in self-reported rates of bullying behavior and victimization have yielded inconsistent results. Nansel and colleagues (2001) found that African American youth were bullied significantly less often than Caucasian or Latino/a youth. Other investigators have reported lower self-reported victimization rates among Latino/a student when compared to all other racial and ethnic groups (Graham & Juvonen, 2002; Hanish & Guerra, 2000). Bullying tends to peak between the ages of 11 and 13 or at about the period of transition from elementary to secondary school (Smith, Madsen, & Moody, 1999).
Researchers and practitioners speak generally of bullying as a phenomenon that involves complex interactions between bullies, victims, bully-victims, and bystanders. Longitudinal studies of bullying suggest that bullies and victims are at elevated risk for later mental health problems and involvement in antisocial conduct (Hawker & Boulton, 2000; Olweus, 1993; Roland, 2002). Both bullies and victims experience elevated levels of depression during adolescence (Forero, McLellan, Rissel, & Bauman, 1999). Bullies participate in other types of aggressive, deviant, and criminal behavior at higher rates than victims (Sourander, Helstela, Helenius, & Piha, 2000), while victims are more likely than bullies to report feelings of insecurity, low self-esteem, and anxiety (Bond, Carlin, Thomas, Rubin, & Patton, 2001; Forero et al., 1999). Victimization is associated with poor academic performance in several studies (Eisenberg, Neumark-Sztainer, & Perry, 2003; Juvonen, Nishina, & Graham, 2000). Evidence suggests that youth who both bully and are bullied by other students exhibit the poorest psychological and social outcomes (Austin & Joseph, 1996; Juvonen et al., 2003).

**Bullying Prevention Approaches**

Prevention programs have used school-wide (Olweus, 1993), classroom management (Roland & Galloway, 2002), peer support (Cowie & Wallace, 2000), playground aggression reduction (Cunningham, Cunningham, Martorelli, Tran, Young, & Zacherias, 1998; Frey et al., 2005) and cognitive-behavioral (Jenson & Dieterich, 2007) strategies to reduce bullying behavior and victimization. Olweus (1993) is generally credited with developing the first efficacious anti-bullying program. The Olweus Bullying Prevention Program was created in response to several youth suicides in Norway during the early-1980s in which bullying appeared to play a major cause. The program is a school-wide approach that has been tested using age-cohort designs in Norway. The strategy raises attention to school bullying by involving students, teachers, and
other school personnel in efforts to change the climate and values associated with bullying (Limber, 2004).

Farrington and Ttofi (in press) recently reviewed the efficacy of 30 bullying prevention programs and found that bullying behavior among students in schools with anti-bullying interventions was approximately 23 percent less than control group schools. Of the 30 programs reviewed by Farrington and Ttofi (in press), 12 interventions were considered effective, 10 were labeled as probably effective, and eight yielded no positive effects on bullying or victimization. Only nine of the 30 studies reviewed were randomized experiments. Parent training, playground supervision, school conferences, disciplinary practices, videos, and classroom management approaches were associated with reductions in bullying. Program elements that emphasized cooperative group work, skill training, peer interventions, parent training, and multimedia strategies were associated with a decrease in victimization. In general, school-wide interventions targeting school and classroom norms about aggression had a greater effect on reducing bullying (e.g., Olweus, 2004) while skill training and other individual-focused interventions had a greater effect on reducing victimization (e.g., Jenson & Dieterich, 2007).

The limited evidence to support or repute the efficacy of anti-bullying interventions in public schools argues for additional intervention research aimed at preventing or reducing bullying and victimization. In the present study, we report outcomes from a group-randomized trial assessing the effects of a skills-based curriculum targeting bullying and bully victimization among elementary school students in a large, urban school district. The Youth Matters (YM) program was tested under controlled conditions in 28 public elementary schools in Denver, Colorado. We present time 4 results detailing the effects of the curriculum on self-reported bullying and bully victimization of fourth- and fifth-grade students. Program and methodological
Aggression and Bullying

lessons learned from the YM trial are discussed and implications for school-based intervention research are noted.

The Denver Public Schools Youth Matters Prevention Trial

Intervention Elements

The YM program promotes the healthy development of young people by encouraging positive relationships between students and school adults and promoting safe and healthy norms throughout the school community. The curriculum consists of a series of instructional modules that address issues (topics) and skills important to students and their school community. Each module includes a 30-40 page story that is designed to help schools meet academic standards in both health education and language arts. The substantive content of each story is directly linked to skills taught in the curriculum modules. To address systemic issues associated with bullying, modules conclude with the development of classroom or school-wide projects that demonstrate the adverse consequences of bullying and aggression to students.

The YM curriculum aims to enhance student’s emotional regulation skills in an attempt to prevent or respond appropriately to bullying. As shown in Figure 1, skill modules emphasize social competency and social resistance skills (e.g., asking for help, making better choices, standing up for yourself and others, preventing bullying behaviors, coping with bullying, etc.) that students can use to stay out of trouble, build positive relationships, make good decisions, and avoid antisocial behavior. In issue modules, students discuss critical developmental concerns (e.g., being a good friend, teasing and bullying, building empathy, risks and norms surrounding aggression, etc.) and create projects that promote positive norms in school. Instructional materials in the issue modules are intended to strengthen peer and school norms against
antisocial behaviors. Understanding the consequences of bullying from both a bully and victim perspective is emphasized.

Four YM curriculum modules were implemented and tested in grades four and five over the course of two academic years. Grades four and five were selected based on an appropriate fit between developmental ability and curricula content and on evidence suggesting that bullying tends to peak in the upper grades of elementary school (Nansel et al., 2001). In addition, grade five marks the transition year prior to middle school enrollment in the district and was identified by school administrators and teachers as a pivotal intervention period for students. Curriculum modules were translated into Spanish for use in the three Spanish speaking classrooms included in the investigation.

Theoretical Foundation

The YM curriculum is guided by theoretical constructs outlined in the social development model [SDM] (Catalano & Hawkins, 1996). The SDM integrates perspectives from social control theory (Hirschi, 1969), social learning theory (Bandura, 1989), and differential association theory (Sutherland, 1973; Matsueda, 1982). The theory proposes that four factors inhibit the development of antisocial behaviors in children: 1) bonding, defined as attachment and commitment to family, school and positive peers (Garmezy, 1985); 2) belief in the shared values or norms of these social units; 3) external constraints such as clear, consistent standards against antisocial behavior (Hansen, Malotte, & Fileding, 1988; Scheier & Botvin, 1998); and 4) social, cognitive and emotional skills that provide protective tools for children to solve problems (Rutter, 1985), assertively and confidently perform in social situations (Werner and Smith, 1982), and resist influences and impulses to violate their norms for behavior (Hansen, Graham,
Aggression and Bullying

Sobel, Shelton, Flay, & Johnson, 1987). Curriculum sessions addressed each of these four core areas. Additional information about the SDM is found in Catalano and Hawkins (1996).

Procedures

The sampling frame for the study consisted of fourth-grade classrooms at 40 Denver elementary schools. Schools were selected by the school district on the basis of risk criteria that included expulsion rates, suspension rates, and the percentage of students eligible for the free lunch program. A decision was made to limit the initial sampling frame to 40 of the 72 elementary schools in order to target those schools with the highest rates of problem behavior. The 40 schools were stratified by geographic region in the city and risk criteria and then randomly assigned to either the control or experimental condition. Fourteen of the 20 schools in each study condition agreed to participate resulting in a total of 28 schools.

Participants

We attempted to consent 674 eligible students in the control schools. Sixty-eight percent of youth in the control group schools consented, 16 percent declined, and 16 percent did not return the consent. Thus, at baseline the control group contained 462 consented students. In the experimental schools we attempted to consent 928 eligible students. Of those attempts, 76 percent consented, 11 percent declined, and 14 percent did not return the consent. The experimental group contained 702 consented students at baseline. Forty-nine percent (N = 328) of students in the experimental group were female compared to 53 percent (N = 242) in the control group. Males accounted for 51 percent (N = 342) and 47 percent (N = 214) of subjects in the experimental and control groups respectively. Sixty-five percent (N = 433) of subjects in the experimental group were Latino/a, 13 percent (N = 88) were African American, 14 percent (N = 94) were American Indian, Asian American, or mixed race/ethnicity, and 8 percent (N = 55)
were Caucasian. Fifty-one percent ($N = 233$) of subjects in the control group were Latino/a, 21 percent ($N = 95$) were American Indian, Asian American, or mixed race/ethnicity, 17 percent ($N = 78$) were African-American, and 11 percent ($N = 50$) were Caucasian. Students in the sample study were similar to other students in the school district with regard to gender and ethnicity.

**Measures**

Student data were collected through classroom surveys conducted in the fall and spring semesters of each academic year. Study outcomes used in the current analysis were bully victimization and bullying behavior. Bully victimization was measured by the bully victim scale from the Revised Olweus Bully/Victim Questionnaire (Olweus, 1996). The scale items are prefaced by the statement “Have you been bullied since the beginning of the school year in one or more of the following ways?” The scale consists of the following items: (1) I was called mean names, made fun of or teased in a hurtful way, (2) Other students kept me out of things on purpose, excluded me from their group of friends or completely ignored me, (3) I was hit, kicked, pushed, or shoved around, (4) Other students spread false rumors about me and tried to make others dislike me, (5) I had money or other things taken from me or damaged, and (6) I was threatened or forced to do things I didn’t want to do. The items were coded as follows: $1 = \text{It hasn’t happened since the beginning of the school year}$, $2 = \text{Only once or twice}$, $3 = \text{2 or 3 times a month}$, $4 = \text{About once a week}$, and $5 = \text{Several times a week}$. The follow-up version of this scale is identical except that the time frame is defined as “in the past month.” A bully victim scale score was calculated by summing the items and dividing by the number of non-missing items, using a threshold of a minimum of four non-missing items to obtain a valid scale score. The Cronbach’s alpha coefficient for this scale is .81.
The bullying behavior outcome was measured by the bullying other students scale from the Olweus Questionnaire. The bullying other students scale contains a set of bullying items that are phrased from the point of view of bullies (e.g., “I called another student mean names, made fun of them or teased them . . .”). The alpha for the bullying other students scale is .80.

Analytic Approach

The data structure consists of four measurement occasions with 1,126 students in 132 classrooms within 28 schools. To test the effects of the program we fit multilevel linear growth models to the study outcomes. Because students changed classrooms but not schools in year two of the study the levels are not completely nested. Thus, the appropriate analytic approach for these data is a cross-classified multilevel model. Measurement occasion is at level 1, and level 2 is a cross-classification of classrooms by students that is nested within schools at level 3. The cross-classified multilevel model that was fitted to the continuous measure of bully victimization is depicted below using the notation of Rasbash and Browne (2001).

\[
y_{i(j_1,j_2)k} = (XB)_{i(j_1,j_2)k} + \nu_k + u_{j_1k} + u_{j_2k} + e_{i(j_1,j_2)k} \tag{1}
\]

Here, \( y_{i(j_1,j_2)k} \) is the scale score at occasion \( i \) in the cross-classification cell defined by student \( j_1 \) in classroom \( j_2 \) in school \( k \). The term \( \nu_k \) is the random effect for school \( k \); \( u_{j_1k} \) is the random effect for student \( j_1 \) in school \( k \); \( u_{j_2k} \) is the random effect for classroom \( j_2 \) in school \( k \); and \( e_{i(j_1,j_2)k} \) is the level 1 residual at measurement occasion \( i \) in the cross-classification cell defined by student \( j_1 \) in classroom \( j_2 \) in school \( k \). The term \((XB)_{i(j_1,j_2)k} \) stands for the linear predictor at occasion \( i \) for student \( j_1 \) in classroom \( j_2 \) in school \( k \).
For each study outcome, the analysis began with a model that included main effects for time, gender, ethnicity, age, and school size plus interaction effects for time by gender and time by ethnicity. At this stage, random effects for the intercept and slope of time were tested by comparing models with and without constraints on the variance components. Next, curvilinear trends were tested by adding quadratic and then cubic terms for time. Models were compared using the likelihood ratio test. Finally, a main effect term for the intervention and an interaction term for time by intervention were added in two steps. At each step effects were assessed using the likelihood ratio test. To obtain estimates of simple main effects and simple slopes for the time by intervention interaction, separate models were fitted with time centered at baseline and time centered at the end of the study, respectively. We followed standard recommended procedures for assessing the adequacy of the fitted models and the tenability of model assumptions (Raudenbush & Bryk, 2002; Singer & Willett, 2003; Snijders & Bosker, 1999).

Selected Findings

Demographic and baseline characteristics. The estimation sample for the current analyses consists of 456 control students and 670 experimental students who were consented at baseline and have at least one data point. Table 1 shows the distribution of demographic characteristics in the intervention and control groups. There is a significantly higher proportion of Latino/a students in the intervention group compared to the control group, $\chi^2 (df=1) = 15.70, \ p < .001$. There is a significantly higher proportion of other ethnic group students in the control group relative to the intervention group $\chi^2 (df=1) = 5.56, \ p < .01$. There were no significant proportional differences across the experimental conditions for the remaining ethnic group categories or gender. Finally, there was no significant difference in the mean age at baseline between the intervention and control groups.
Between 10 and 15 percent of all male participants indicated they had called other students mean names, excluded students from peer-group activities, or hit, kicked, and shoved other classmates at baseline. Rates of these same bullying behaviors for girls ranged from four to eight percent. Student reports of victimization at baseline were considerably higher than self-reports of bullying behavior. Between 25 and 30 percent of boys indicated they had been the victim of name calling, exclusion, physical aggression, and rumor spreading. Girls reported similar rates of victimization for all behaviors except physical aggression; only 20 percent of female participants said they had been the recipient of physical bullying at baseline.

**Intervention effects.** We did not find a significant difference in rates of bullying between subjects in the experimental and control group schools over time. Bullying behaviors decreased in both groups; the decline was greater in YM schools but was only marginally significant ($p < .10$) when compared to control group schools at the end of year two (Jenson & Dieterich, 2007). We did, however, find significant effects in reducing bully victimization among subjects in the experimental group.

Table 2 shows the results of fitting a linear growth model to the continuous measure of bully victimization. In comparison to the baseline model that included main effects and interactions for a linear term for time, gender, ethnicity, age, and school size, a model including a quadratic term for time fit the data better ($\chi^2 = 12.96$, df=1, $p < .001$). Addition of the main effect for intervention to the curvilinear trend model did not improve model fit ($\chi^2 = 1.08$, df=1, p-value=.299). However, addition of the intervention by time interaction term did improve model fit ($\chi^2 = 4.36$, df=1, p-value=.037).
Estimates of the simple slopes and simple main effects associated with the time interaction terms are displayed in Table 2. The coefficients for time reflect the instantaneous rates of change in bully victimization at baseline and end of study, respectively, for a Caucasian male student in the control group. Note that the instantaneous time slope is negative at baseline and positive by the end of the study. The coefficient for time squared is the constant decelerating effect on the negative slope of time. The coefficients for YM indicate that the two conditions did not differ on the bully victim scale at baseline, but by the end of the study, there is a significant intervention effect. The coefficient for the interaction of time by YM indicates that bully victimization is declining at a faster rate in the intervention condition relative to the control condition. Bully victimization is declining at a faster rate among African American students compared to Caucasian students. African American and Caucasian students report similar levels of bully victimization at baseline, but by the end of the study African American students report significantly lower levels of bully victimization.

Figure 2 compares the predicted trajectories from the continuous outcome growth model of bully victimization for an average intervention school and an average control school adjusting for gender, ethnicity, age, and school size. The trajectories reveal significant intervention effects on the rate of change for the bully victim scale. Results indicate that self-reported bully victimization among students in the YM schools decreased at a significantly higher rate compared to students in control group schools. By the end of the study, bullying victimization decreased in the YM schools by approximately 20 percent compared to nine percent in the control group schools.
Practice and Methodological Lessons from the *Youth Matters* Trial

Implications for School-Based Bullying Prevention Programs

Reductions in bully victimization found in the Denver YM trial are encouraging in view of the fact that curriculum modules focused on teaching the requisite social and emotional skills necessary to deal effectively with the threat of bullying. It may be that the acquisition of skills aimed at discouraging and coping with bullying incidents had a positive effect on reducing victimization reports by students in the YM schools. Interestingly, reductions in bully victimization found in the YM trial are consistent with findings from previous investigations in which skills-based curricula were used as the primary intervention (Farrington & Ttofi, in press).

Surprisingly, outcomes from the YM trial reveal no significant differences in rates of bullying between subjects in experimental and control group schools at the end of the intervention. Bullying declined in both groups, and while the decline in bullying was greater among subjects in the YM schools, the difference in self-reported bullying between students in the experimental and control groups was only marginally significant at the end of the study. One point to consider with regard to this finding is the possibility that students may actually occupy the status of bully and victim at the same time. Further, it is possible that students may also switch their status between committing bullying behaviors and experiencing bully victimization over time. Thus, it may be that the growth modeling technique used in our analyses is not the best means to assess change in bully and victim status across time points.

We recently conducted a Latent Transition Analysis (LTA) to better understand the factors that influence changes in bully status over the course of time. Findings from the LTA yielded three distinct latent classes of victims, bully-victims, and bystanders. Interestingly, no
distinct bullying group was found suggesting that children in the Denver trial who reported bullying other students were also the victims of bullying (Powell, Jenson, & Dieterich, 2009). This finding, while contrary to other studies in which separate bully groups have been identified, may be explained by the frequent peer rejection and adverse reputations that many bullies report experiencing at school (Giang & Graham, 2008; Gifford-Smith & Brownell, 2003). It also suggests that childhood bullies may be far less powerful or influential than students and teachers perceive them to be. Regardless, this finding challenges the common assumption that bullies are a separate group that is somehow immune from bully victimization at school.

Finally, results from the YM trial point to the need for anti-bullying interventions that include both school-wide and individually-focused components. That is, to be effective, bullying prevention programs should include strategies aimed at promoting positive school norms and individual-level approaches that equip students with the requisite skills necessary to resist or deflect bullying. To date, no single intervention has fully integrated these two alternate intervention strategies.

Methodological Implications for School-Based Intervention Research

The YM trial in Denver used a group-randomized design to examine intervention effects. A group-randomized trial (GRT) was viewed as optimal because it provided a hierarchical or nested design option that accounted for classroom and school effects on bullying and victimization outcomes. Conducting a GRT requires random assignment to an experimental or control condition. Because universal prevention programs like YM are usually delivered to intact schools or classrooms, treatment conditions are defined in terms of groups (classrooms or schools). In the YM study we randomized at the school level due to program feasibility and to
protect against spillover between experimental and control conditions that can occur when randomization takes place at the classroom level (Cook, 2005; Flay & Collins, 2005).

Commensurate with decisions about how best to randomly assign groups are methodological issues pertaining to analytic approach, sampling and statistical power, measurement, and missing data. Recommendations to enhance these and other design elements in school-based intervention research are discussed below.

*Choose an analytic approach appropriate for GRTs.* GRTs provide obvious advantages in nested designs. However, they also create special analytic concerns. Analyzing data from GRTs requires an approach that accounts for the clustering of observations within classrooms and/or schools (Fraser, Richman, Galinsky, & Day, 2009; Murray, Varnell, & Blitstein, 2004). Thus, statistical methods known variously as multilevel, hierarchical, or mixed models are necessary to analyze data from GRTs. These models partition random and fixed effects at the student-, classroom-, and school-level (Raudenbush & Bryk, 2002).

*Develop a sampling plan that meets the analytic requirements of GRTs.* Randomization at the school level means that the number of groups in nearly all GRTs is relatively small. The small number of groups (schools) creates distinct sampling and analytic challenges because sample size is determined by the number of groups randomized, not by the number of students. Various design elements influence the number of groups required for adequate power in GRTs. Chief among these elements are the size and variability of the treatment effect and the strength of the intraclass correlation (ICC). The ICC is the proportion of the total explainable variance that occurs at the group level (Raudenbush & Bryk, 2002). It is commonly viewed as a measure of the degree of dependence of observations within the same group or cluster (Snijders & Bosker, 1999). Although sample size requirements for statistical power depend on many factors, usually
20 to 30 groups are required in a GRT. Practically speaking a study of 20 to 30 schools is a large-scale undertaking, but statistically, a sample of 20 to 30 schools is quite small. Finally, randomizing units from samples this small increases the risk of bias arising from unequal distribution of nuisance variables (Murray, 1998). Thus, in the YM trial we stratified schools by four geographic quadrants in Denver to mitigate potential racial or ethnic bias across study conditions.

Select measures and design longitudinal data collection protocols. Measures in GRTs, like in all intervention studies, must be linked to key components of the intervention and be sensitive to change. Classroom surveys are the most common method of assessing program outcomes in school-based prevention. Self-report protocols are typically read aloud to students as they fill out individual answers at their desks. We used two time points per academic year in our study, one each during fall and spring semesters. Our measures assessed risk and protective factors for problem behavior and participation in victimization, bullying, and other antisocial conduct.

Design the follow-up period and data collection schedule to optimize validity and precision. Data collection over three or more time points requires more sophisticated analyses of change than the simple pretest and posttest designs that are often used to evaluate school-based interventions (Singer & Willett, 2003). As described earlier, multilevel linear growth models were used in the YM trial. One advantage of a linear growth model is that it can handle unbalanced or missing data as long as certain assumptions are met, as opposed to standard repeated measures approaches that require balanced and complete data for subjects across all data points. An introduction to modeling change in multilevel data is available in Singer and Willett (2003).
Anticipate and plan for missing data. Another methodological problem peculiar to school-based research that warrants mention is missing data resulting from student mobility and attrition. Although multilevel models can handle incomplete data, mobility and attrition can pose a threat to validity if they are substantial enough. Excessive missing data can also cause numerical and convergence problems in model development. Thus, mobility and attendance patterns in the school district should be assessed during the design phase of a study. Depending on the type and severity of mobility and attrition, multiple membership models (Rasbash, Steele, Browne, & Prosser, 2004) or crossed-effects models (Raudenbush & Bryk, 2002) such as the one used in the YM study may need to be fitted. These models may also be required for the analysis of data from longitudinal studies that follow students through the transitions from elementary to middle school or middle school to high school.

Critical Issues in School-Based Intervention Research

Intervention research in public schools raises a number of practical issues that are likely to be common challenges for investigators. A brief set of recommendations follows.

Elicit school district support early in the process. School district prevention, research, and planning offices are important places to begin dialogue about conducting intervention research. In the YM study, district administrators representing each of these offices were able to help us refine our research design based on the characteristics and organizational structure of the school district. They also provided key supports that enabled the research team to inform teachers and principals about the benefits of using the YM curriculum.

Target known risk and protective factors for problem behaviors. Prevention curricula are most effective when they target risk and protective factors associated with problem behavior (Jenson & Fraser, 2006). In many cases data gathered from routine classroom and school-level
reports may be adequate to identify and target salient risk and protective factors across schools. Once factors are identified, prevention efforts should be tailored to target the specific factors that are most salient for students. For example, school district and community data may indicate that risk factors of frequent residential mobility and school failure and protective factors of positive attachments to adults or extended family members occur frequently in a specific geographic region. It follows, then, that prevention efforts in local schools would target specifically these factors in curricula design and implementation. Strategies consistent with these risk and protective factors might seek to enhance educational supports such as tutoring, systematically track student mobility each semester, and develop a mentoring component.

Consider age and developmental phase. The question of intervention timing has been discussed extensively in the prevention literature. Gottfredson and Wilson (2003) found that most school-based prevention programs are aimed at early adolescent and middle (junior high) school students. This may be because problem behaviors such as aggression, drug use, and delinquency begin to occur at higher rates during the early adolescent years. School personnel and research teams should, therefore, consider the prevalence of problem behaviors and the placement of current prevention programs when considering at which age to target prevention activities. The YM investigation targeted upper elementary school children in an effort to reduce the onset of bullying and aggression just prior to middle school. This decision was based on local and national data indicating that bullying and bully victimization often peak in grades four and five.

Develop and implement interactive curricula. Effective prevention programs use interactive techniques such as peer discussion groups and cognitive-behavioral training strategies to engage students (Gottfredson & Wilson, 2003; Tobler, 2000). Interactive skill training
techniques are associated with positive outcomes in a number of prevention experiments. Social, behavioral, and cognitive skills are among the most frequently taught skills in universal and selected prevention programs (Catalano, Berglund, Ryan, Lonczak, & Hawkins, 2002).

**Reinforce core academic competencies.** Emphasis on core academic competencies in reading, language, and math has increased in recent years. Therefore, intervention researchers must recognize the importance of integrating prevention activities and academic content. The YM curriculum does not explicitly teach core competencies. However, curriculum content and activities include reading, story prediction, and reflective journaling that aim to support overall reading skills. Clearly, prevention programs like YM will have to attend carefully to academic content in lesson design as the emphasis on academic skills and standardized testing in schools increases (National Institute on Drug Abuse, 2003).

**Consider program length and duration.** Evidence examining the relationship between program length and outcomes is somewhat mixed. Gottfredson and Wilson (2003) examined 94 prevention programs and found that program length was not related to outcomes. However, in an earlier review, Tobler and Stratton (1997) noted that higher intensity programs characterized by at least 16 hours were more effective than low intensity strategies (6 hours or less). Further, Catalano et al. (2002) found that prevention programs that lasted nine or more months were more effective than interventions lasting less than one academic year. Finally, evidence from longitudinal studies suggests that interventions introduced and taught over the course of several years are more likely to yield positive outcomes when compared to single-shot interventions of relatively brief duration (e.g., Scheier, Botvin, Diaz, & Griffin, 1999). In consideration of these factors, the YM trial tested the effects of four curriculum modules with the same students over two academic years.
Be sensitive to gender, race and ethnicity. Strategies to best meet the diverse needs of racial and ethnic groups in prevention programs have been the topic of considerable discussion. While evidence suggests that effective prevention programs work with a broad range of students (e.g., Botvin, Griffin, Diaz, Ifill-Williams, 2001), tailoring interventions to specific populations does increase efficacy for racial and ethnic subgroups (Botvin, Schinke, Epstein, Diaz, & Botvin, 1995; Castro & Alarcon, 2002). Based on such evidence, the most common approach to ensuring relevance to different racial and ethnic groups is to adapt or modify an existing curriculum to meet the needs of particular groups (Ringwalt, Vincus, Ennett, Johnson, & Rohrbach, 2004).

Given the fact that cultural adaptation is common, the task for investigators and the field becomes one of how to accomplish adaptation while maintaining program fidelity. Several program adaptation models have been identified. Backer (2001) describes an adaptation guide that aims to strike a balance between fidelity and implementation. The model includes 12 sequential steps to follow in the adaptation process; evidence from studies using the specified model has yet to be extensively reported. Castro, Barrera, and Martinez (2004) note that cultural program adaptations must go beyond simple changes in structure such as altering the appearance of role models to targeting the core beliefs and values associated with different populations. They emphasize that adaptation must address modifications in both program content and delivery. Addressing cultural relevancy in content and program delivery is consistent with what Castro et al. (2004) identify as hybrid models that aim to bridge cultural relevancy and fidelity. Martinez and Urbana (2002) developed such a model in their work with Hispanic families at the Oregon Social Learning Center and we followed this model of adaptation in the Denver trial.

Establish rigorous training, implementation, and fidelity protocols. The position, selection, and training of the individuals who deliver prevention content are important moderator
variables in prevention programs. Many programs train classroom teachers to deliver intervention materials. One major advantage of this approach is that classroom teachers are in an optimal position to reinforce aspects of the curriculum during all hours of a typical school day. Alternatively, programs may hire and train their own curriculum specialists to deliver the intervention. Outside trainers are often in a better position than classroom teachers to ensure fidelity of intervention content. However, because their time in the classroom is limited specialists are at a disadvantage when it comes to reinforcing elements of the curriculum during other times of the day. We used outside curriculum specialists to deliver the intervention in the YM trial. Our experience suggests that strong classroom management skills and experience working with groups of children were essential to delivering the curriculum.

Random fidelity checks to monitor the implementation of curricula used in school-based intervention trials are common in school-based prevention research. Fidelity has been the topic of numerous papers and considerable discussion in the field of prevention (e.g., Elliott & Mihalic, 2004). Experts suggest that fidelity measures include observations of duration, sequence, and the degree to which instruction complies with overall curriculum objectives (Ringwalt, Ennett, Vincus, Thorne, Rohrbach, & Simons-Rudolph, 2002). In our study, the project director used a structured checklist to indicate the extent to which curriculum activities are followed by trainers. The results of these sessions were then reviewed with curriculum specialists to improve subsequent classroom lessons.

In sum, school-based intervention research requires attention to a myriad of program, methodological, and practical concerns and issues. Addressing these issues requires a team of investigators and clinicians who can navigate successfully the many political and research
challenges inherent in intervention research. In the following section we discuss ways to increase school-based intervention research by social work investigators.

Intervention Research in Social Work

Intervention research is at a critical juncture in social work. Prior studies have shown that relatively few social work investigators conduct intervention research and that social work lags significantly behind intervention research in disciplines such as public health and psychology (e.g., Jenson, in press). Further, a very small percentage of peer-reviewed publications in the social work literature describe outcomes from intervention trials (Rosen, Proctor, & Staudt, 1999). The lack of attention to intervention research in social work is somewhat surprising in view of past and current efforts aimed at integrating elements of practice (intervention) and science in social work education and research (Briar, 1974; Fraser et al., 2009; Kirk & Reid, 2002; Rothman & Thomas, 1994).

How can we increase the volume and quality of intervention research in social work? We offer three suggestions to enhance intervention research.

*Teach principles of intervention research across the educational continuum.* Clearly, at its highest level, intervention research is a complicated endeavor that requires a host of substantive, political, and methodological skills. However, the core steps in conducting basic intervention research – specifying a problem, creating or selecting an intervention or program, and testing effectiveness – are in many ways similar to the steps used by experienced clinical and community practitioners. Exposure to the process and discussion of the importance of intervention research should be increased at all levels of social work education. Undergraduate students should be taught to recognize the core principles of accountability that are evident in the underlying values of intervention research. MSW-level training should be redesigned to include
practice and research courses devoted to improving intervention development and research skills. Doctoral students should learn and apply the requisite methodological and analytical training that is necessary to conduct all phases of intervention research. Students should also be encouraged to apply intervention research methodology to a substantive area.

*Link elements of intervention research to evidence-based practice and translational research.* Current attention to evidence-based practice (EBP) and translational research offers a unique opportunity to advance intervention research in social work. The original definition of EBP outlines a five-step process that requires practitioners to ask relevant questions and locate evidence that is necessary to select and evaluate appropriate interventions for clients (Gambrill, 2006; Sackett, Rosenberg, Gray, Haynes, & Richardson, 1997). The process of EBP in many ways parallels the basic steps of intervention research. Similarities in the two concepts should be promoted in education and practice settings.

A second emphasis of EBP has been directed at increasing the awareness of empirically-supported interventions among practitioners, administrators, and policy officials (Woolf, 2008). We have learned a great deal about the effectiveness of interventions for a broad range of individual and social problems in the past several decades. Importantly, research, policy, and governmental entities concerned with improving the dissemination of efficacious interventions have created registries of effective programs that are widely available to practitioners, educators, and the general public (e.g., Campbell Collaboration Library, 2008; Center for the Study and Prevention of Violence, 2008; Schinke, Brounstein, & Garnder, 2002). Such dissemination efforts, viewed by many as key products or outcomes of EBP, should be effectively conveyed to the practice community.
The larger issue of translating knowledge gained from intervention research to the field is a challenging one. To date, translational research mechanisms in social work have proceeded at a rather slow pace (Brekke, Ell, & Palinkas, 2007). Thus, concerted efforts among social work education, community agencies, and policy organizations are needed to bring translation and dissemination of efficacious interventions to the field.

One example of effective translation is evident in the dissemination efforts associated with cost-benefit analyses of prevention programs being conducted by Aos and colleagues (Aos et al., 2004). Results of these analyses are providing important and practical data to policymakers and practitioners as they search for solutions to complex funding and program implementation questions. A second example of effective translation is seen in the current test of the Communities That Care (CTC) prevention approach (Hawkins, Oesterle, Brown, Arthur, Abbot, Fagan, & Catalano, 2009). CTC is a prevention system that guides community stakeholders through a structured process in which they identify prevalent risk and protective factors for problem behavior in their towns or cities. Stakeholders and coalitions are then encouraged to choose tested and effective programs to implement in their communities. Results from a randomized trial of CTC involving 24 small towns in seven states indicate that eighth-grade youth in the experimental (CTC) condition report significantly less alcohol, binge drinking, smokeless tobacco, and delinquency than youth in control group communities (Hawkins et al, 2009). CTC holds great promise as an effective way to help communities adopt tested and effective programs to preventing childhood and adolescent problems.

Adopt successful strategies from the fields of public health and prevention science. Public health principles have been widely used in the past decade to develop and test interventions in schools (Catalano, 2007; Woolf, 2008). Outcomes from controlled prevention trials indicate that
empirically based and theoretically sound prevention curricula can prevent the onset of child and adolescent problems (Wilson, Lipsey, & Derzon, 2003). To illustrate, recent meta-analyses and systematic reviews have identified a number of efficacious interventions aimed at preventing aggression and violence (Hahn et al., 2007; Limbos, Chan, Warf, Schneir, Iverson, Shekelle, & Kipke, 2007; Wilson & Lipsey, 2007) drug use (Foxcroft, Ireland, Lister-Sharp, Lowe, & Breen, 2003) and delinquency (Catalano, Arthur, Hawkins, Berglund, & Olson, 1998; Catalano, Loeber, & McKinney, 1999).

These developments led to the establishment of what is now commonly called prevention science (Coie, Watt, West, Hawkins, Asarnow, Markman, Ramey, Shure, & Long, 1992; Hawkins, Catalano, & Arthur, 2002). Based on public health principles, the key elements of prevention science imply that: 1) factors associated with a problem behavior must be changed in order to prevent that behavior; 2) malleable risk and protective factors identified in empirical studies should be the targets of prevention efforts; 3) prevention programs should be rigorously tested; and 4) efficacious programs should be disseminated and implemented with fidelity in school and community settings. Knowledge gained from studies guided by a public health approach to prevention has significantly improved our understanding of the onset and persistence of childhood and adolescent problems (Biglan, Brennan, Foster, & Holden, 2004). We believe that social work research would benefit greatly by following the standards of intervention research evident in public health and prevention science.

Summary

Significant advances in school-based intervention research have been realized in the past two decades. Results from studies such as the Denver YM trial indicate that well-designed
programs can prevent or reduce childhood and adolescent problem behaviors. Furthermore, many investigations have demonstrated that efficacious interventions are cost-effective.

Social work faces tremendous challenges in its efforts to promote intervention research. Few schools of social work use principles of intervention research as a teaching or curricula framework. Training, mentoring, and institutional barriers to producing skilled intervention researchers must be identified and eliminated. Papers in this special issue identify a number of ways to promote and increase intervention research. These suggestions should be taken seriously by educators, researchers, and practitioners at all levels of influence.
References


Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Prevention & Treatment, 5*, Article 15.


Table 1. Demographic characteristics by study condition

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>Intervention</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants (n)</td>
<td>456</td>
<td>670</td>
<td>1126</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>9.79</td>
<td>9.84</td>
<td>9.82</td>
</tr>
<tr>
<td>SD</td>
<td>.48</td>
<td>.51</td>
<td>.50</td>
</tr>
<tr>
<td>Gender (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>53.1</td>
<td>49.0</td>
<td>50.6</td>
</tr>
<tr>
<td>Ethnicity (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latino/a</td>
<td>52.9</td>
<td>64.6</td>
<td>59.9</td>
</tr>
<tr>
<td>African American</td>
<td>16.7</td>
<td>13.1</td>
<td>14.6</td>
</tr>
<tr>
<td>Anglo</td>
<td>11.2</td>
<td>8.2</td>
<td>9.4</td>
</tr>
<tr>
<td>American Indian</td>
<td>3.3</td>
<td>3.1</td>
<td>3.2</td>
</tr>
<tr>
<td>Asian</td>
<td>3.1</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>.0</td>
<td>.4</td>
<td>.3</td>
</tr>
<tr>
<td>Other</td>
<td>12.9</td>
<td>7.3</td>
<td>9.6</td>
</tr>
</tbody>
</table>
Table 2. Results from the continuous outcome growth model of bully victimization with baseline and end-of-study centering of the time variable

<table>
<thead>
<tr>
<th></th>
<th>Baseline-centered</th>
<th>End-of-study-centered</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Intercept)</td>
<td>-.062 (.098)</td>
<td>-.024 (.111)</td>
<td>.533</td>
</tr>
<tr>
<td>Time Squared</td>
<td>.021 (.007)</td>
<td>.021 (.007)</td>
<td>.001</td>
</tr>
<tr>
<td>Time</td>
<td>-.078 (.041)</td>
<td>.097 (.042)</td>
<td>.065</td>
</tr>
<tr>
<td>Latino</td>
<td>-.051 (.085)</td>
<td>-.202 (.102)</td>
<td>.547</td>
</tr>
<tr>
<td>Black</td>
<td>.023 (.100)</td>
<td>-.270 (.120)</td>
<td>.821</td>
</tr>
<tr>
<td>Other</td>
<td>.090 (.096)</td>
<td>-.127 (.114)</td>
<td>.346</td>
</tr>
<tr>
<td>Female</td>
<td>.015 (.046)</td>
<td>.072 (.055)</td>
<td>.748</td>
</tr>
<tr>
<td>Age_c</td>
<td>-.008 (.040)</td>
<td>-.008 (.040)</td>
<td>.836</td>
</tr>
<tr>
<td>Size50_c</td>
<td>-.012 (.010)</td>
<td>-.012 (.010)</td>
<td>.224</td>
</tr>
<tr>
<td>YM</td>
<td>.015 (.078)</td>
<td>-.171 (.083)</td>
<td>.852</td>
</tr>
<tr>
<td>Time:Latino</td>
<td>-.037 (.030)</td>
<td>-.037 (.030)</td>
<td>.213</td>
</tr>
<tr>
<td>Time:Black</td>
<td>-.071 (.035)</td>
<td>-.071 (.035)</td>
<td>.040</td>
</tr>
<tr>
<td>Time:Other</td>
<td>-.053 (.034)</td>
<td>-.053 (.034)</td>
<td>.117</td>
</tr>
<tr>
<td>Time:Female</td>
<td>.014 (.016)</td>
<td>.014 (.016)</td>
<td>.392</td>
</tr>
<tr>
<td>Time:YM</td>
<td>-.045 (.022)</td>
<td>-.045 (.022)</td>
<td>.045</td>
</tr>
</tbody>
</table>

Variance Components

<table>
<thead>
<tr>
<th></th>
<th>Student</th>
<th>Class</th>
<th>School</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>intercept</td>
<td>.326</td>
<td>&lt;.001</td>
<td>.016</td>
</tr>
<tr>
<td></td>
<td>time slope</td>
<td>.014</td>
<td>&lt;.001</td>
<td>.016</td>
</tr>
</tbody>
</table>

Notes: For the test of the YM effects, the degrees of freedom are calculated as $J - S - 1$, where $J$ is the number of groups (28) and $S$ is the number of school-level predictors on either the intercept ($S=2$) or the slope of time ($S=1$). The critical $t$-value for the two-tailed test with 25 degrees of freedom is 2.06. Each variance component was tested using a likelihood ratio test that compared the restricted versus the unrestricted model.
Figure 1. *Youth Matters* intervention model

- Bullying Threats or Incidents
- Effective Emotional Regulation
- Social, Behavioral, and Cognitive Skills
- Positive Classroom Norms
- Reduction in Bullying

*YM Components*
- Skills training
- Classroom activities
- Reading and journaling
- School-wide projects
Figure 2. Bully victimization: Predicted curves for control and intervention schools adjusted for ethnicity, gender, age, and school size

\[ \beta_{YM} = -0.171, \quad t\text{-value} = -2.074 \]