

2009

## Observed Reductions in School Bullying, Nonbullying Aggression, and Destructive Bystander Behavior: A Longitudinal Evaluation

Karin S. Frey

Miriam K. Mirschstein

Leihua V. Edstrom

Jennie L. Snell

Follow this and additional works at: [https://digitalcommons.georgefox.edu/gscp\\_fac](https://digitalcommons.georgefox.edu/gscp_fac)

 Part of the [Psychology Commons](#)

---

# Observed Reductions in School Bullying, Nonbullying Aggression, and Destructive Bystander Behavior: A Longitudinal Evaluation

Karin S. Frey and Miriam K. Hirschstein  
University of Washington

Leihua V. Edstrom  
Bellevue School District, Bellevue, Washington

Jennie L. Snell  
University of Washington

This study was a longitudinal extension of a random control trial of the Steps to Respect antibullying program. Students in Grades 3–5 were surveyed ( $n = 624$ ) and observed on the playground ( $n = 360$ ). Growth curve models of intervention students showed 2-year declines in playground bullying, victimization, nonbullying aggression, destructive bystander, and argumentative behavior. Grade-equivalent contrasts indicated group differences in all problem behaviors. Problem behaviors in the control group increased or remained stable across grade. Intervention group students reported less difficulty responding assertively to bullying compared with control students. Within both groups, older students perceived themselves to be more aggressive and less frequently victimized than younger students. Methodological issues posed by inconsistencies between self-reported and observed behavior are discussed.

*Keywords:* bullying, victimization, bystanders, intervention, observations

School bullying is a widespread problem that compromises opportunities to learn, disrupts the social connections among students, and detracts from the positive quality of classroom experiences. Observations show an average of 2.4 bullying episodes occurring every hour within classrooms (Craig, Pepler, & Atlas, 2000). Whether bullying happens inside or outside of the classroom, teachers must contend with the impact on children's social, emotional, and academic adjustment. Class participation declines among students who are excluded by peers, and school avoidance increases among children who are victims of overt bullying (Buhs, Ladd, & Herald, 2006). Victimization predicts poor academic performance both concurrently and up to a year later (Juvonen,

Nishina, & Graham, 2000; Nishina, Juvonen, & Witkow, 2005; Schwartz, Gorman, Nakamoto, & Toblin, 2005). Links between victimization and school achievement and attendance appear to be mediated by physical symptoms and psychological distress associated with victimization (Nishina et al., 2005). Furthermore, changes in victimization and distress predict subsequent school adjustment, including achievement, attendance, and teacher-rated cooperativeness (Juvonen et al., 2000).

Young people who frequently bully also perform poorly academically (Nansel et al., 2001), although this finding may be most typical of students using overt methods of coercion (cf. Schwartz et al., 2005). Bullying predicts date abuse (Connolly, Pepler, Craig, & Taradash, 2000), substance abuse (Pepler, Craig, Connolly, & Henderson, 2002), and street violence (Andershed, Kerr, & Stattin, 2001), conditions that may also affect interactions on school grounds.

Witnessing bullying is distressing for bystanders (Charach, Pepler, & Ziegler, 1995) and may contribute to greater approval for aggressive retaliation (Musher-Eizenman et al., 2004) and an increase in weapon carrying among victimized children (DeVoe & Kaffenberger, 2005). Fear of bullying is so pervasive in some schools that students are reluctant to seek help from adults (Unnever & Cornell, 2003, 2004). Such cultures of bullying may seriously undermine the ability of educators to mentor students and provide a safe learning environment.

Bullying is an intentional aggressive process that involves using unequal power to cause harm or secure material and social benefit. It occurs within social–ecological contexts that support coercive behavior (Olweus, 1993; Pellegrini & Long, 2002; Pepler, Craig, & O'Connell, 1999; Swearer & Espelage, 2004), conditions that are often met on school grounds. Supervising adults rarely intervene (Craig & Pepler, 1997; Frey et al., 2005). Aggressive and nonaggressive classmates also play key roles either impeding or

---

Karin S. Frey, Educational Psychology, College of Education, University of Washington; Miriam K. Hirschstein, School of Nursing, University of Washington; Leihua V. Edstrom, Bellevue School District, Bellevue, Washington; Jennie L. Snell, Department of Psychology, University of Washington.

Data collection and some analyses for the study were completed while we were employed by Committee for Children, Seattle, Washington, who provided the funding. Currently, none of us have any financial affiliation with the grantor. The views expressed herein are ours and have not been cleared by the grantor.

A portion of the data was previously reported in the *Proceedings of Persistently Safe Schools 2005: The National Conference of the Hamilton Fish Institute*. We thank the school personnel, the children, and their parents for consenting to test the program. We are grateful to Diana Lim for contributing the study graphs. Diane Jones and Sue Nolen provided cogent reviews of an earlier version of the manuscript. We are indebted to Debra Pepler and Wendy Craig for allowing us to review their videotapes of playground aggression.

Correspondence concerning this article should be addressed to Karin S. Frey, Educational Psychology, College of Education, Box 3600, University of Washington, Seattle, WA 98195. E-mail: karinf@u.washington.edu

encouraging aggression (Espelage, Holt, & Henkel, 2003; Farmer et al., 2002; Pellegrini, Bartini, & Brooks, 1999; Rodkin, Farmer, Pearl, & Van Acker, 2006; Salmivalli, 1999). Although bystanders' rare efforts to terminate bullying are usually successful (Craig et al., 2000), bystanders more often reward bullies with increased respect, attention, and active encouragement (Craig & Pepler, 1997; Salmivalli, 1999). Compared with other forms of adolescent aggression, bullying shows greater variation across cultures (Smith-Khuri et al., 2004), providing further evidence of the importance of social-ecological context.

Nonbullying aggression may also contribute to a climate that supports bullying. Victims who respond aggressively to their tormentors appear to suffer continued and even increased bullying (Kochenderfer-Ladd, 2003; Salmivalli & Helteenvuori, 2007). Compared with children who bully, are not involved, or respond passively to bullying, aggressive victims tend to have difficulty managing emotions and respond with hostility to provocations (Nansel et al., 2001; Schwartz, 2000; Toblin, Schwartz, Gorman, & Abou-ezzeddine, 2005). Whereas some bullying can be executed so skillfully that it is difficult to identify the perpetrator (Garandeau & Cillessen, 2005), hostile or reactive aggression is highly visible and associated with peer rejection (Card & Little, 2006; Coie & Dodge, 1998). The lack of support from peers (Nansel et al., 2001) and even adults (Frey & Nolen, in press) may make reactively aggressive children attractive targets.

Like aggressive victims, passive victims also show skills deficits, such as the ability to respond assertively on their own behalf (Toblin et al., 2005). Although perpetrators and supervising adults bear the ultimate responsibility for peer abuse, concerted efforts to foster social-emotional skill may help children protect themselves and elicit support from others.

Because social behavior is multiply determined, programs may be most successful if they adopt multiple intervention strategies. Frey and Nolen (in press) described a model of school-based intervention based on social-ecological models of bullying (Olweus, 1993; Pepler et al., 1999; Swearer & Espelage, 2004), transactional analyses of intervention processes (Sameroff & MacKenzie, 2003), and social-cognitive models of development (Bandura, 1986; Coie & Dodge, 1998; Huesmann, 1988). The social ecology identifies the appropriate context for intervention, whereas transactional and social-cognitive models identify the processes by which ecosystems change or remain static. In a successful intervention, social interactions reflect changes that have occurred throughout the community and that stimulate additional changes in the thought, emotions, and behavior of individuals. To adequately address risk factors for school bullying, environmental interventions should (a) increase adult knowledge and intervention, (b) minimize rewards for coercive behavior, (c) forge socially responsible peer norms (e.g., Henry et al., 2000; Olweus, 1993; Swearer & Espelage, 2004), and (d) support efforts to improve social-emotional and behavior skills (Schwartz, 2000; Toblin et al., 2005).

Much of the thinking about bullying intervention has been inspired by Olweus's (1993) innovative schoolwide program that showed large reductions in bullying and victimization among Norwegian students. More recent evaluations of bullying interventions have been somewhat disappointing (e.g., Jenson & Dietrich, 2007). A meta-analysis (Smith, Schneider, Smith, & Ananiadou, 2004) showed small or even negative effects of schoolwide pro-

grams. Smith and colleagues (2004) attributed the results to several possible factors, including cultural differences and inadequate implementation.

Most evaluations have relied almost exclusively on student self-reports of victimization and bullying. These have poor to low correspondence with peer reports, teacher reports, or objective observations (Cornell & Brockenbrough, 2004; Pellegrini & Bartini, 2000) and yield grade trends in victimization that diverge from other sources (Salmivalli, 2002). Investigators have raised concerns that exposure to intervention materials may bias posttest reporting (Schafer, Werner, & Crick, 2002) by sensitizing students to the bullying they perpetrate and experience.

The use of real-time observations has been an underutilized tool in the evaluation of school bullying, especially considering the important information such observations have provided about the environmental context and contributions of peers to the maintenance and cessation of bullying (e.g., Craig & Pepler, 1997; Hanish, Ryan, Martin, & Fabes, 2005). Without observations, most evaluations have been methodologically constrained by the difficulty of discriminating between changed perceptions measured via self-reports and actual behavior change. Objective evidence showing that schoolwide programs reduce bullying and victimization is therefore limited. Snyder et al. (2006) argued that observations are particularly well suited to intervention research because of blinding to intervention status and sensitivity to behavior change. Another benefit is that trained observers appear to differentiate between reactive aggression and the instrumental aggression typical of bullying better than other reporters (Card & Little, 2006). The current study makes use of both objective playground observations and subjective reports.

### The Steps to Respect Program

Bullying, and the rewards thereof, increase at the end of elementary school (Hanish & Guerra, 2004; Pepler et al., 2002). Acceptance of aggression also increases and begins to stabilize then (Huesmann & Guerra, 1997). Because behavior, beliefs, and patterns of interaction that are longstanding become increasingly resistant to change (Coie & Dodge, 1998; Huesmann & Guerra, 1997), the Steps to Respect program (Committee for Children, 1997) targets the upper elementary school years as a particularly important time to influence bullying-related skills, beliefs, and behavior. The program is designed to decrease school bullying problems by (a) increasing adult monitoring and intervention in bullying events, (b) improving systemic supports for socially responsible behavior, (c) changing student normative beliefs that support bullying, and (d) addressing student social-emotional skills that counter bullying and support social competence. Steps to Respect is a multilevel program that coordinates a schoolwide environmental intervention, three sequential classroom curricula, and a selected intervention for students involved in bullying. The program includes manuals, written material, and audiovisual presentations for school administration, staff training, classroom curricula, and parent outreach. Professional staff training and implementation consultation are available.

### *Schoolwide Environmental Intervention*

In preparation for classroom lessons, school administrators and staff establish schoolwide bullying policies and procedures that

protect reporting students and encourage socially responsible actions. Disciplinary models encourage proportional, consistent actions aimed at stopping problems before they escalate. Other preparatory actions include clarification of adult responsibilities, identification of areas needing increased supervision, and procedures to ensure appropriate follow up. Staff training and written materials provide models that can be adapted for the specific school infrastructure.

A core instructional session for all school staff aims to dispel myths suggesting that bullying is inconsequential or usually perpetrated by easily identified “problem” students. Staff receive an overview of program goals and key features of program content. Teachers, counselors, and administrators receive additional training in how to coach students involved in bullying episodes. Third-through sixth-grade teachers are provided with a curriculum orientation. They practice role-playing exercises, an often omitted aspect of classroom lessons (Kallestad & Olweus, 2003).

### *Classroom Curricula*

Classroom lessons and practices use cognitive-behavioral techniques to promote socially responsible norms (Huesmann & Guerra, 1997) and to foster social-emotional skills. Lessons and instructional practices for the third through sixth grades include 10 weeks of biweekly basic lessons followed by 8 to 10 literature-based lessons.

Because normative beliefs about aggression stabilize around fourth grade (Huesmann & Guerra, 1997), lessons place bullying within a moral context (Terasahjo & Salmivalli, 2003). They counter beliefs that bullying is harmless or that personal characteristics of victims justify bullying (Rigby & Slee, 1991). The program also aims to counter expectations that students can bully with impunity. Prescriptions for socially responsible behavior emphasize the avoidance of destructive bystander behavior and the provision of support for targeted classmates. Reporting bullying is defined as “telling to keep someone safe,” whereas tattling is defined as “telling to get someone in trouble.” Copies of policies are sent to parents and reiterated in class posters, parent nights, and assemblies.

Steps to Respect curricula provide multiple pathways to influence social behavior by building bullying prevention and general social-emotional skills. Bullying prevention skills include assertive responding, appropriate help seeking, and problem solving that emphasizes safety and avoidance of future problems. Social-emotional skills include perspective taking, emotion management, conflict resolution, and other skills that promote caring relationships (Espelage & Swearer, 2003). Because friendships can provide a buffer to bullying (Boulton, Trueman, Chau, Whitehand, & Amatya, 1999), lessons provide opportunities to practice friendship skills.

### *Individual Intervention*

Brief individual coaching sessions with each participant in bullying episodes are intended to provide solution-oriented responses to immediate and long-term student needs. Coaching protocols (one for perpetrators, one for targets of bullying) provide strategies to establish facts, empower students to avoid future problems, and assess effectiveness. While not ignoring the need for sanctions,

coaching sessions focus on empathy, problem solving, and practicing assertiveness skills. Teachers discuss school norms and collective responsibility for school safety (see Frey, Edstrom, & Hirschstein, in press, for more details on program content).

### Previous Evaluation and Current Hypotheses

A random control evaluation of the Steps to Respect program using objective observers showed promising changes in playground bullying, destructive bystander behavior, and argumentative and agreeable social behavior (Frey et al., 2005). It may, however, take several years for a school to consolidate and fully implement the antibullying policies, procedures, and classroom practices associated with comprehensive programs. Some behaviors may also require more extensive intervention than others before showing change. After 6 months of intervention, intervention group students involved in fall bullying or destructive bystander behavior differed significantly from their control group counterparts (Frey et al., 2005). Nonbullying aggression and victimization showed no improvement at the group level. Only intervention teachers who intervened more frequently with individuals involved in bullying episodes saw victimization and nonbullying aggression decrease in their classrooms (Hirschstein, Edstrom, Frey, Snell, & MacKenzie, 2007). The current study extends the first year evaluation by examining student behaviors, beliefs, and self-reported behavior following a second year of implementation. Data from students in the intervention group (Frey et al., 2005) were collected for a second year and compared with that of control group students of equivalent grades. This enabled us to examine longitudinal hypotheses regarding program effects on students’ playground behaviors, beliefs, and self-reported behaviors and to address questions raised by first year findings.

### Hypotheses: Combined Effects of Level 1 and Level 2 Lessons

#### *Playground Behaviors*

We expected that results of the intervention would become stronger as the program continued through its second year. We predicted greater reductions in observed bullying behavior and expected to find reductions in nonbullying aggression. Previously observed reductions in argumentative behavior and increases in agreeable behavior among boys (Frey et al., 2005) were expected to continue and strengthen.

First year comparisons used pretest occurrence or nonoccurrence of problem behaviors to determine whether group differences were due to prevention effects (e.g., reducing the growth of bullying among pretest nonperpetrators) and/or intervention effects (e.g., reducing bullying among pretest perpetrators). Both bullying and encouragement of bullying showed significant intervention effects (Frey et al., 2005). In this study, we continued to look for evidence of both intervention and prevention effects by examining longitudinal change among those initially involved or uninvolved in bullying.

#### *Student Beliefs and Self-Perceptions*

At the 6-month posttest, fifth- and sixth-grade intervention group (but not control group) students reported increased confi-

dence in their ability to respond assertively to bullying (Frey et al., 2005). At the 18-month posttest, the longitudinal sample students were in the fourth and fifth grades. We expected to see increases in confidence following implementation of the second program level.

In the first year analyses, group differences in normative beliefs occurred as a result of deterioration in the control group rather than improvements in the intervention group (Frey et al., 2005). That is, students in control but not intervention schools became more accepting of bullying and aggression, viewed adults as increasingly unresponsive to bullying problems, and felt less responsibility to report bullying or persuade friends to desist. This is consistent with previous studies that show increasing acceptance of aggression (Huesmann & Guerra, 1997; Pellegrini & Long, 2002) and decreasing sympathy for victims (Rigby & Slee, 1991) across this age range. In light of this evidence and ceiling effects in pretest responses, we speculated that preventing deterioration during the second year was a more likely outcome than was improvement.

### Self-Reported Behaviors

In a similar vein, developmental trends and initial basement effects in self-reported direct and indirect aggression may have masked group differences at the 6-month posttest. Victimization showed a near-significant reduction in the intervention group relative to the control group (Frey et al., 2005). Hypotheses regarding self-report measures were complicated by concern that the intervention might increase the salience of bullying and victimization in a way that masked group differences. Reasoning that sensitization effects of the program may have lessened by the second year, we looked for group differences, with intervention students reporting lower levels of direct aggression, indirect aggression, and victimization.

## Method

### Design

To foster study participation and school commitment after random assignment to condition, control schools received Steps to

Respect training and materials after 1 year of study participation. Outcomes from the control year were used to provide time-lagged contrasts between grade-equivalent groups (Olweus, Limber, & Mihalic, 1999).

The third- and fourth-grade intervention students (from the sample in Frey et al., 2005) comprised the longitudinal sample for the current study. These students received the curricula in order: Level 1 in the first intervention year and Level 2 in the second. Assigning two grade levels to the longitudinal sequence enabled us to increase the sample size as well as examine possible grade differences in effectiveness. The current study examined longitudinal patterns in bullying-related behavior and attitudes over 18 months. To provide context for intraindividual change, we compared the longitudinal sample with control students in the same grade. The sequence of activities for students in the longitudinal intervention and control groups is presented in Table 1.

### Participants

Six elementary schools in the Pacific Northwest participated in the study. Schools within two suburban districts were matched for size, ethnic breakdown, and percentage of students receiving free and reduced lunch (range = 21% to 60%). To facilitate playground observations of a relatively large sample, we staggered school participation. A matched pair in one district started participating in the 2000–2001 school year. Two matched pairs of schools in another district were added in the 2001–2002 school year. By spring of 2002, 2 years of observations were completed for the longitudinal sample (now fourth and fifth grades) of the first cohort intervention school, as well as the control school observations for the second cohort and first year observations for the remainder of the longitudinal sample (two schools). The longitudinal sample of the second cohort was completed in the spring of 2003. The three control schools received the intervention after 1 year of study participation.

*Students.* Active parental consent was obtained for 64% of students in the third through sixth grades, yielding an original sample of 1,126 students. Child assent was obtained from fourth to sixth graders during survey administration.

Table 1  
Data Collection and Program Implementation Schedule for Each Cohort

Group	Student participation					
	Year 1			Year 2		
	Pretest <sup>a</sup> grade	Classroom curricula	6-month posttest <sup>a</sup> grade	12-month posttest <sup>b</sup> grade	Classroom curricula	18-month posttest <sup>a</sup> grade
Intervention	3	Level 1	3	4	Level 2	4
	4	Level 1	4	5	Level 2	5
Control <sup>c</sup>	3		3			
	4		4			
	5		5			

<sup>a</sup> Teacher ratings, student surveys, and observations collected. <sup>b</sup> Reported data only. <sup>c</sup> Data collection for the control schools coincided with data collection for the first year data in some intervention schools and the second year data in other intervention schools.

Third- and fourth-grade students in the intervention group were followed for 2 years. Of the 225 students, 95 were assessed in the third grade and again in the fourth grade, and 130 were assessed in the fourth and fifth grades. The delayed-intervention control group consisted of 399 students (132 in third grade, 123 in fourth grade, and 144 in fifth grade) who were tested in the year prior to receiving the intervention. A subset of 164 longitudinal students and 196 control group students were randomly selected at pretest for playground observation. The longitudinal and control samples used in the current study were equally divided by gender (49.4% female), which did not vary by group.

Student ethnic background and English proficiency were reported by teachers. There were no intervention–control group differences in ethnic makeup (Frey et al., 2005), and the ethnic makeup did not change from the first to the second year (8.6% African American, 14.5% Asian American, 5.5% Hispanic American, 1.4% Native American, and 70.0% European American). Similarly, the proportion of students speaking English as a second language did not vary by condition and remained unchanged in the second year of study participation (11.5%).

*Teachers.* The design required that fourth-grade teachers, who had taught Level 1 in the first intervention year, teach Level 2 in the second. Of the 22 intervention teachers in the longitudinal sample, 8 teachers continued to teach fourth grade. Two fourth-grade teachers were new to the program (and received training in the program in the second year). Eight fifth-grade teachers delivered the same Level 2 lessons that they had in the first year. One fifth-grade teacher was new to the program, 2 had previously taught Level 1, and 1 had previously taught Level 3. There were 30 control group teachers in the third through fifth grades. All teachers agreed to complete study measures, for which they received monetary compensation. Experimental teachers also agreed to periodic observation of program lessons. None had previous experience with the Steps to Respect program.

## Implementation

*Implementation sequence.* Table 1 shows the implementation and data collection schedule. In the first year of study participation, school teams in the intervention group developed antibullying policies and systems to handle bullying reports in September and October. Committee for Children trainers conducted standard professional development activities in three after-school sessions. Intervention school personnel were trained in November. Classroom lessons were implemented from December or January through May. Playground observations were not collected in the fall of the second year, enabling us to allow greater flexibility in the initiation and pacing of the lessons (per teacher request). In most classes, lessons began in the fall and ended between December and March. Thus, the span of time between lesson completion and posttest observations was somewhat longer in the second year of study participation than in the first.

*Implementation fidelity.* Lesson dosage and fidelity were similar in the first and second year of study participation. In the second year, all but 2 teachers reported completing all 10 skill lessons. Those 2 reported teaching 7 of 10. Eighteen of the 21 teachers completed a language arts unit designed to reinforce lessons presented in the skill lessons.

Observers documented lesson adherence and quality. Interrater agreement was measured on 61 occasions. Computation of weighted kappas found  $\kappa = .81$  for completion of lesson objectives and  $\kappa = .62$  for lesson quality. Because of funding issues, there were only three lesson observations for the two schools in the second cohort, compared with five in the first. We computed average scores on the basis of all available observations. During the observations, teachers completed 91% of the lesson objectives and received a mean lesson quality rating of 2.24 (on a scale from 1 = *poor* to 3 = *good*). These did not vary from ratings in the first year.

## Study Procedure

Playground observations were collected between October and December (pretest) and between April and June (6-month posttest) of the first intervention year and again during April to June of the second intervention year (18-month posttest). Each child was observed for 5-min sessions approximately once a week over the two 10-week periods. Mean observation times were 149 min. Only children meeting the a priori minimum of 40 min of observation during each observation period were included in longitudinal analyses at the 18-month posttest. Comparisons between the longitudinal and control samples included students who had 40 min of observation time for the first and last time points. Common reasons for incomplete data included a child moving to another school or missing recess.

## Playground Observations

*Development of the coding system.* The coding system was created following multiple live and videotaped observations of playground behavior by Jennie L. Snell and Karin S. Frey. These observations were used to translate conceptual distinctions between bullying and nonbullying aggression into operational definitions. Within 5-min coding periods, coders could be expected to recognize power differentials when a larger child was intimidating another or when a group of children were obviously whispering and snickering about a nearby child. For more subtle distinctions, we relied on previous research indicating that most aggression occurring between asymmetrically aggressive dyads is proactive, with one child repeatedly targeted for aggression (Dodge, Price, Coie, & Christopoulos, 1990). Although bullying is not limited to proactive aggression, a substantial percentage of proactive aggression is bullying (Coie & Dodge, 1998). Therefore, when coders observed two instances within 5 min in which a particular child behaved aggressively toward a specific child who did not reciprocate, they coded both instances as bullying. The coding system was constructed so that aggression was identified as bullying at the end of the bout or, in the case of nonreciprocated aggression, later in the session. This enabled coders to collect the necessary information to reliably differentiate the two types of aggression.

The coding manual (Snell & Frey, 2000) was refined during the first 2 months of observer training conducted by Jennie L. Snell. Data collection revealed that some behaviors occurred too infrequently on playgrounds to obtain interrater reliability. These were included in a more general category (e.g., peer support for a victim is positive social behavior) or, in the case of teacher intervention in a bullying event, dropped from analyses entirely.

*Observational procedures and codes.* Observers were recruited through newspaper advertisements, class announcements, and word of mouth. Most of those selected for training were recent college graduates in psychology. One was a former teacher. Two were graduate students with flexible schedules. To be able to complete observations, we staggered the start dates into two waves. Six observers were hired for the first wave of data collection in two schools. Fourteen observers collected data the next year. They observed at the four new schools in the fall and spring. Spring observations also included the longitudinal sample (fourth and fifth grades) from the first wave intervention school. Longitudinal observations for the second wave intervention schools were completed in the third year by returning observers.

For 12 weeks, observers spent 15 hr a week memorizing the codes and viewing and coding playground videotapes. Prior to progressing to live coding on playgrounds, coders were required to demonstrate agreement with a master coder with a minimum of  $\kappa > .70$  (Cohen's kappa) on a series of increasingly difficult training tapes. One trainee was not employed beyond the training period because of failure to reach criterion. All posttest observers coded during the pretest, though not necessarily at the same schools. Prior to posttest observations, we held a 1-month refresher training. The observers needed to demonstrate coding proficiency on a new set of videotapes ( $\kappa > .70$ ) before advancing to playground observations.

To reduce child reactivity, we preceded data collection at each time point by a 2-week habituation period on the playground. During these periods, observers needed to reach the criterion of  $\kappa > .70$  on playground coding before starting data collection. The order of children observed was randomly determined on a daily basis, with the proviso that observers would proceed to the next child on the list if the preceding one was not available. Five-minute continuous focal-child samples were collected using handheld computers with a custom-designed program. Sessions that were shorter than 5 min (because of recess ending or child going to the office) were supplemented with additional coding periods. The goal was to obtain a minimum of 10 observation sessions over the 12-week period. Mean observation times for the pretest, 6-month posttest, and 18-month posttest were 49.6 min, 49.7 min, and 49.3 min, respectively. Random agreement checks were made by master coders throughout data collection to prevent decay. Overall agreement was calculated for 15% of the sessions ( $n = 210$ ) on an event-by-event basis in single observation sessions (overall  $\kappa = .80$ ). To provide a more stringent test of reliability, we also calculated separate agreement statistics for each code (provided below).

Actual kappas for the analyzed data are likely to be higher than those reported below because they were aggregated over 10 sessions (Hubbard, Dodge, Cillessen, Coie, & Schwartz, 2001). A final analysis to examine the ability of coders to distinguish bullying from nonbullying aggression indicated excellent discriminability ( $\kappa = .80$ ).

1. Bullying ( $\kappa = .63$ ): physical, verbal, or indirect aggression involving either (a) a discernible power imbalance between aggressor(s) and target (e.g., a group of children aggressing against a single child) or (b) repeated aggression, during the same observation session, by a child toward a nonretaliating peer.

2. Nonbullying aggression ( $\kappa = .54$ ): physical, verbal, or indirect aggression that did not involve a discernible power imbalance or repeated nonreciprocal aggression. Bossy or argumentative behavior was not coded as aggression. Bullying and aggression were distinguished from mock fighting or playful teasing by the absence of mutual expressions of pleasure or interest.
3. Destructive bystander behavior ( $\kappa = .55$ ): laughter or cheers during bullying events or sustained passive watching while an aggressive act took place within 15 feet (4.572 m) of the focal child.
4. Argumentative social behavior ( $\kappa = .56$ ): nonaggressive, negative acts or statements directed toward another child. This category included behaviors such as acting bossy, arguing, or ignoring another child's attempts to enter a group.
5. Agreeable social behavior ( $\kappa = .81$ ): neutral or positive acts or statements toward another (e.g., starting a conversation).

### *Teacher Ratings and Student Surveys*

Student survey data and teacher ratings of student interpersonal skills were collected over a 2-week period in the fall (pretest and 12-month posttest) and spring (6- and 18-month posttests) of each intervention year. This yielded four data points for the longitudinal sample.

In October and April, teachers rated students with the 17-item Peer-Preferred Social Behavior subscale of the Walker-McConnell Scale of Social Competence and School Adjustment, Elementary Version (Walker & McConnell, 1995). Calculated at each data point, reliability coefficients (alphas) ranged from .85 to .89.

At the same time points, students completed the School Experience Survey (Frey et al., 2005). Scales included Acceptance of Bullying/Aggression (e.g., "It's okay to say something mean to a kid who really makes you angry;" seven items,  $\alpha$ s ranged from .86 to .88), Bystander Responsibility (e.g., "If my friends were telling lies about another kid I would tell them to stop;" five items,  $\alpha$ s ranged from .86 to .89), Perceived Adult Responsiveness (e.g., "Adults at my school stop kids from being bullied;" four items,  $\alpha$ s ranged from .59 to .66), Difficulty of Responding Assertively (e.g., "Kids at school are teasing you. How hard would it be to calmly tell them to stop?;" five items,  $\alpha$ s ranged from .80 to .84), Direct Bullying/Aggression (e.g., "I called kids names at school;" eight items,  $\alpha$ s ranged from .87 to .89), Indirect Bullying/Aggression (e.g., "I told my friends to ignore kids that I was mad at;" six items,  $\alpha$ s ranged from .76 to .82), and Victimization (e.g., "A group of kids at school called me mean names;" eight items,  $\alpha$ s ranged from .83 to .86).

## Results

We first present data pertaining to student retention rates over 18 months and group comparisons of pretest variables. Second, we provide the results of three-level mixed hierarchical models (SPSS 14.0) for observed behavior, student beliefs, self-reported aggres-

sion and victimization, and teacher-reported peer interaction skill. The models nested time point (fixed effect, Level 1) within individual student (random effect, Level 2). Covariates for gender and grade were entered as fixed Level 2 variables and retained in the model if significant as main effects or in interactions. These, in turn, were nested within the second year classroom (random effect, Level 3), thus controlling for shared error among classmates (Hedeker, Gibbons, & Flay, 1994).

Finally, we used three-level modeling to provide grade-equivalent comparisons between the longitudinal intervention and the control samples. Values assigned to grade indicated whether the data point corresponded to the fall (e.g., 3.0, 4.0) or the spring (e.g., 3.5, 4.5) testing period.

### Student Retention and Pretest Group Comparisons

Retention rates from the pretest to the 6-month posttest were 91.3% for the larger group of intervention and control group students and 92.4% for the observation subsample. Rates did not differ by group. After 2 years, retention rates were 74.0% for the larger longitudinal sample and 81.3% for the observation subsample. Because we were comparing 2 years of intervention group data with 1 year of control group data, we were concerned that lower retention in the longitudinal sample might yield a select group with fewer behavior problems than in the control sample. Comparisons of pretest data for the two groups, however, revealed no differences in observed behaviors (all  $F$ s < 1). Self-reported behavior and attitudes also revealed no group differences, with  $F$  values that exceeded 1 only for direct aggression,  $F(1, 419) = 2.19$ ,  $ns$ , and perceived bystander responsibility,  $F(1, 419) = 2.70$ ,  $ns$ .

### Longitudinal Analyses of Playground Behaviors in Intervention Schools

Three-level mixed models (Singer & Willet, 2003) were used to examine the intervention group students at three time periods (pretest, 6-month posttest, and 18-month posttest). Our analyses divided students into those who exhibited specific problem behaviors at pretest and those who did not. Pretest involvement was used as one of two fixed factors in the design, 3 (time period)  $\times$  2 (pretest involvement), with time period analyzed at Level 1 and

pretest involvement at Level 2. Student characteristics (e.g., sex, grade) and interactions with time period were also included in the model if significant (Level 2). To address problems of skew and kurtosis in the raw data, we subjected frequencies for all observation codes except agreeable behavior to log transformations. Simple effects for pretest involvement were found for all playground codes and are not discussed further. Results of the fixed effects tests for log-transformed observed behaviors are presented in Table 2, along with untransformed rates for each behavior at the three time periods. Effect sizes (Cohen's  $d$ ) were also calculated to assess the net change for program participants. Means and standard deviations from the pretest and 18-month posttest were used to avoid inflated values because of the correlation between the time points (Dunlap, Cortina, Vaslow, & Burke, 1996).

**Bullying behaviors and victimization.** Bullying in intervention schools showed a significant decline across the three time periods. As shown by the Time  $\times$  Pretest Involvement interaction (see Table 3), the significant overall decline in bullying is attributable to changes in the behavior of those who bullied during the pretest period ( $p < .01$ ,  $d = 2.11$ ).

Victimization by bullying consistently declined over the three time periods, with no overlapping confidence intervals. As shown in Table 3, this pattern was directly attributable to declines in victimization among those who were victimized at pretest ( $p < .01$ ,  $d = 1.24$ ).

A significant effect of time period for destructive bystander behavior was similarly modified by a Time  $\times$  Pretest Involvement interaction. Confidence intervals show consistent declines among those who previously encouraged bullying, particularly after the first year of program implementation ( $p < .01$ ,  $d = 2.26$ ).

For each of these three bullying-related behaviors, significant differences between those involved and not involved at pretest (all  $F$ s > 196.27,  $p < .01$ ) disappeared by the 18-month posttest (all  $F$ s < 1). The longitudinal patterns for those involved in problem behaviors at pretest show means declining to the low levels of those who were not involved at pretest. The levels for those not involved at pretest showed nonsignificant increments and may reflect random measurement error.

**Other playground behavior.** Nonbullying aggression in intervention schools showed significant declines modified by a significant

Table 2  
Three-Level Mixed Models of Intervention Group Playground Behavior at Three Time Points

Observed behavior	Type III tests of fixed effects		Time period means (and standard deviations)			
	Time period	Time period $\times$ Pretest involvement	Pretest	6-month posttest	18-month posttest	Cohen's $d$
Rates per hour						
Bullying	$F(2, 436) = 15.28^{**}$	$F(2, 437) = 72.62^{**}$	0.67 (1.14)	0.54 (1.23)	0.49 (0.91)	.17
Victimization	$F(2, 436) = 22.14^{**}$	$F(2, 436) = 78.57^{**}$	0.57 (1.15)	0.46 (0.92)	0.36 (0.84)	.21
Bystander	$F(2, 432) = 74.27^{**}$	$F(2, 433) = 125.19^{**}$	0.54 (0.95)	0.32 (0.74)	0.15 (0.47)	.63
Nonbullying aggression	$F(2, 338) = 7.77^{**}$	$F(2, 339) = 51.3^{**}$	1.40 (1.92)	1.12 (2.25)	0.89 (1.54)	.29
Percentage of time						
Argumentative	$F(2, 333) = 18.61^{**}$	$F(2, 339) = 26.04^{**}$	3.25 (3.01)	2.70 (2.68)	2.20 (2.59)	.37
Agreeable	$F(2, 310) = 2.48^{\dagger}$	$F(2, 309) = 15.14^{**}$	41.01 (15.02)	43.11 (16.68)	40.60 (15.54)	0

Note. For the Type III tests of fixed effects, degrees of freedom vary as a function of the magnitude of classroom-level effects. Cohen's  $d$  compares means and standard deviations of the pretest and 18-month posttest.

\*\*  $p < .01$ .  $\dagger p < .10$ .

Table 3

*Estimated Marginal Means of Intervention Group Playground Behavior as a Function of Pretest Involvement and Time Point*

Observed behavior	Students involved at pretest			Students not involved at pretest		
	<i>M (SE)</i>	95% CI	<i>n</i>	<i>M (SE)<sup>b</sup></i>	95% CI <sup>b</sup>	<i>n</i>
	Rates per hour (log transformed)					
Bullying			51			96
Pretest	.44 (.02)	.39, .49		.00		
6-month posttest	.15 (.02)	.11, .20		.11 (.02)	.08, .15	
18-month posttest	.13 (.03)	.08, .18		.11 (.02)	.07, .14	
Victimization			46			101
Pretest	.42 (.02)	.38, .46		.00		
6-month posttest	.17 (.02)	.13, .22		.10 (.02)	.07, .13	
18-month posttest	.08 (.02)	.03, .12		.10 (.02)	.07, .13	
Bystander			43			104
Pretest	.43 (.02)	.40, .47		.00		
6-month posttest	.11 (.02)	.08, .15		.06 (.01)	.04, .08	
18-month posttest	.05 (.02)	.01, .08		.04 (.01)	.01, .06	
Nonbullying aggression			76			71
Pretest	.52 (.03)	.46, .58		.00		
6-month posttest	.27 (.03)	.21, .33		.14 (.03)	.08, .20	
18-month posttest	.22 (.03)	.16, .27		.13 (.03)	.07, .19	
	Students above the median at pretest					
	Students below the median at pretest					
Observed behavior	<i>M (SE)</i>	95% CI	<i>n</i>	<i>M (SE)<sup>b</sup></i>	95% CI <sup>b</sup>	<i>n</i>
	Percentage of time					
Argumentative <sup>a</sup>			71			76
Pretest	.74 (.03)	.68, .81		.30 (.03)	.24, .37	
6-month posttest	.52 (.03)	.46, .59		.37 (.03)	.30, .43	
18-month posttest	.45 (.03)	.38, .52		.31 (.03)	.24, .37	
Agreeable			72			75
Pretest	51.0 (1.8)	47.55, 54.52		31.3 (1.8)	27.71, 34.90	
6-month posttest	48.3 (1.8)	44.80, 51.77		38.2 (1.8)	34.63, 41.82	
18-month posttest	44.1 (1.8)	40.46, 47.70		37.3 (1.8)	33.66, 40.96	

Note. CI = confidence interval.

<sup>a</sup> Log transformed. <sup>b</sup> Values not provided when means equal zero.

Time × Pretest Involvement interaction. As shown in Table 3, aggression in both posttest periods was lower than in the pretest period for the group that exhibited this behavior at pretest ( $p < .01$ ,  $d = 0.83$ ). Argumentative interactions declined overall. This was due solely to the decline among the more argumentative students after the first year intervention ( $p < .01$ ,  $d = 0.85$ ).

Agreeable interactions, contrary to predictions, did not increase over the three time periods in intervention schools. Inspection of the Time × Pretest Involvement interaction indicates that students initially above the mean in agreeable interactions showed mean declines that were nearly equal to the mean increases found among those initially below the means, strongly suggesting no program effect on this variable.

These three behaviors were the only ones on which girls and boys differed. Girls were less aggressive, less argumentative, and more agreeable than boys.

#### *Longitudinal Surveys of Beliefs and Reported Behaviors*

Student beliefs and reported behaviors were measured four times in the longitudinal intervention group (pretest, 6-month posttest, 12-month posttest, and 18-month posttest). The scale

scores for self-reported direct aggression were subjected to square root transformations to remediate kurtosis problems in the untransformed data. The fixed effects tests for time-point differences are presented in Table 4, along with estimated means.

*Student beliefs.* Students who participated in the Steps to Respect program across 2 school years showed no change in Acceptance of Bullying/Aggression across four survey administrations. Girls were less accepting than were boys ( $p < .01$ ). Contrary to predictions, Perceived Bystander Responsibility decreased over time, with significant differences found between pretest and the 12- and 18-month posttests. Perceived Adult Responsiveness regarding bullying problems dropped in the second year. The 12- and 18-month posttests indicated that students perceived adults to be less responsive than at pretest, particularly in the fall of the second year (12-month posttest). In line with predictions, mounting an assertive response to bullying was judged to be less difficult at the 12- and 18-month posttests than at the pretest for children in the intervention group.

*Self-reported behavior and experiences.* As predicted, self-reports of victimization in intervention schools declined across the three time points (see Table 4). Student reports of their own direct aggression showed a significant interaction of grade with survey time,

Table 4

*Fixed Effects and Means for Individual Change in Beliefs and Reported Behavior Within the Intervention Group*

Variable	Fixed effects test	Item means (and standard deviations)			
		Pretest	6-month posttest	12-month posttest	18-month posttest
<b>Belief</b>					
Acceptance of bullying	$F(3, 647) < 1$	1.59 <sup>a</sup> (.64)	1.60 (.69)	1.59 (.71)	1.62 (.72)
Bystander responsibility	$F(3, 632) = 20.08^{**}$	3.60 <sup>a</sup> (.60)	3.53 (.62)	3.43 (.63)	3.29 (.69)
Adult responsiveness	$F(3, 644) = 68.14^{**}$	3.31 <sup>a</sup> (.60)	3.28 (.64)	2.78 (.54)	3.08 (.64)
Difficulty of assertion	$F(3, 652) = 4.20^*$	2.13 <sup>a</sup> (.74)	2.14 (.78)	2.04 (.76)	1.99 (.82)
<b>Self-reported behavior</b>					
Direct aggression	$F(3, 651) = 1.83$	1.31 <sup>b</sup> (.49)	1.31 (.46)	1.40 (.54)	1.46 (.60)
Indirect aggression	$F(3, 655) = 3.53^*$	1.73 <sup>b</sup> (.65)	1.77 (.68)	1.74 (.68)	1.90 (.78)
Victimization	$F(3, 659) = 4.39^*$	2.06 <sup>b</sup> (.77)	1.91 (.77)	1.96 (.75)	1.90 (.79)
Peer interaction skill	$F(3, 694) = 4.67$	3.78 <sup>b</sup> (.88)	3.93 (.85)	3.85 (.86)	3.88 (.80)

*Note.* Self-reported direct aggression was square root transformed for effect estimation. Means are untransformed. Degrees of freedom vary as a function of the magnitude of classroom-level effects.

<sup>a</sup> Item ranges = 1 to 4. <sup>b</sup> Item ranges = 1 to 5.

\*  $p < .05$ . \*\*  $p < .01$ .

$F(3, 651.4) = 2.89, p < .05$ . Contrary to predictions for the intervention group, this indicated an increase in self-reported direct aggression at the 18-month posttest for the students in the fifth grade in the second year of the intervention ( $p < .001$ ). Reports among the fourth-grade longitudinal sample did not vary across time ( $F < 1$ ). Students' reports of indirect aggression increased over time. Girls indicated less direct aggression and more indirect aggression than did boys (both  $ps < .01$ ).

*Teacher-reported peer interaction skill.* Both time and the Grade  $\times$  Time interaction were significant for teacher ratings of peer interaction skills,  $F(3, 694.0) = 4.82, p < .01$ , with fourth-grade teachers rating students' interaction skills higher at the 12- and 18-month posttests than the third-grade teachers of the same students had at pretest and the 6-month posttest ( $p < .05$ ). The only change noted within year and within teacher, however, was found among the fourth- to fifth-grade cohort. Fourth-grade teachers reported increased skill from the pretest to the 6-month posttest ( $p < .05$ ). None of the other three within-year comparisons showed significant increases. Teachers rated the interaction skills of girls more highly than those of boys ( $p < .01$ ).

*Grade-Equivalent Group Comparisons*

Using three-level analyses with student and classroom as random variables (Levels 2 and 3, respectively), we performed time-lagged, grade-equivalent comparisons between the longitudinal intervention group and the control group. Intervention students were nested within the classroom they attended during the 18-month posttest. Grade was calculated in half-year increments. The fixed effects model yielded a 6 (grade)  $\times$  2 (group) design, with grade analyzed at Level 2 and group effects at Level 3.

The estimates of fixed effects for group and the Group  $\times$  Grade interaction are presented in Tables 5 and 6. Follow-up analyses calculated slopes for each group, which are shown in the last two columns of each table. To facilitate interpretation, we present the slopes based on nontransformed data: Rate per hour for observed bullying and aggression and mean rating per item for self-reported aggression. In these analyses, the values for grade 4.5 combine data from the 18-month posttest for the younger cohort with the data from the 6-month posttest for the older cohort. Because the improvements are larger at the 18-month posttest, the slope coef-

Table 5

*Three-Level Model of Intervention–Control Group Differences in Observed Behavior, Controlling for Grade*

Observed behavior	Fixed effect		Grade slope ( $\beta$ )	
	Group	Group $\times$ Grade	Intervention	Control
Bullying	$F(1, 176) = 8.54^{**}$	$F(1, 202) = 12.86^{**}$	0.04	0.35 <sup>**</sup>
Victimization	$F(1, 270) = 11.82^*$	$F(1, 308) = 16.58^{**}$	0.04	0.25 <sup>**</sup>
Bystander	$F(1, 333) = 6.71^{**}$	$F(1, 407) = 11.22^{**}$	-0.17 <sup>**</sup>	0.15
Nonbullying aggression	$F(1, 365) = 2.29$	$F(1, 477) = 5.13^*$	-0.30 <sup>*</sup>	0.21
Argumentative	$F(1, 377) = 2.70$	$F(1, 480) = 4.74^*$	-0.64 <sup>**</sup>	0.06
Agreeable	$F(1, 802) < 1$	$F(1, 802) < 1$	-0.58	-0.22

*Note.* The log-transformed rate was used to estimate fixed effects for all but agreeable behavior. Slopes were estimated using untransformed rates per hour and percentage of time. Degrees of freedom vary as a function of the magnitude of classroom-level effects. The three-level model includes observation point, student, and classroom. Grades ranged from 3.0 (fall) to 5.5 (spring) in half-year increments.

\*  $p < .05$ . \*\*  $p < .01$ .

Table 6  
*Three-Level Model of Intervention–Control Group Differences in Beliefs and Reported Behavior, Controlling for Grade*

Variable	Fixed effect		Grade slope ( $\beta$ )	
	Group	Group $\times$ Grade	Intervention	Control
<b>Belief</b>				
Acceptance of bullying	$F(1, 1427) = 1.15$	$F(1, 1644) = 3.44^\dagger$	0.02	0.20**
Bystander responsibility	$F(1, 1272) < 1$	$F(1, 1658) < 1$	-0.21	-0.28**
Adult responsiveness	$F(1, 1243) = 2.46$	$F(1, 1644) = 3.78^\dagger$	-0.24**	-0.13**
Difficulty of assertion	$F(1, 1092) = 6.45^{**}$	$F(1, 1518) = 8.37^{**}$	-0.13**	0.07
<b>Self-reported behavior</b>				
Direct aggression	$F(1, 1623) < 1$	$F(1, 1670) < 1$	0.10**	0.13**
Indirect aggression	$F(1, 1621) < 1$	$F(1, 1668) = 2.21$	0.06*	0.10*
Victimization	$F(1, 1422) < 1$	$F(1, 1672) < 1$	-0.08	-0.07
Peer interaction skill	$F(1, 1501) = 2.43$	$F(1, 1732) = 2.63$	0.05	0.17

*Note.* Direct aggression was square root transformed for fixed effect modeling. All slopes are based on untransformed item means. Degrees of freedom vary as a function of the magnitude of classroom-level effects. The three-level model included observation point, student, and classroom. Grades ranged from 3.0 (fall) to 5.5 (spring) in half-year increments.

\*  $p < .05$ . \*\*  $p < .01$ .  $^\dagger p < .10$ .

ficients for the intervention group probably underestimate long-term improvements in the intervention group.

*Observed bullying behaviors.* Mean levels of bullying, victimization, and destructive bystander levels were significantly lower in the intervention group than in the control group and interacted significantly with grade (see Table 5 and Figure 1). Overall, slopes indicated no changes in bullying and victimization from third to fifth grade in the intervention group, whereas both behaviors increased within the control group. Slopes indicated a decline in destructive bystander behavior within the intervention group; the behavior did not change in the control group. Figure 1 shows separate lines for each of the intervention and control samples.

*Other playground behavior.* Significant Group  $\times$  Grade interactions were found for nonbullying aggression and argumentative behavior. Slopes indicated grade-related declines in both behaviors within the intervention group but no change in the control group. As in the longitudinal analyses, agreeable behavior did not conform to predictions and showed no group differences.

*Student beliefs.* As shown in Table 6 and Figure 2, children in the intervention group tended to be less accepting of bullying and aggression than those in the control group ( $p < .10$ ). Visual inspection of the slopes for each group suggests that program participation may at least retard grade-related increases in acceptance of aggressive behavior.

No group differences were found in children's perceived bystander responsibility, which declined over time. Perceptions of adult responsiveness to bullying problems tended to differ by group ( $p < .10$ ) because of the steep drop reported by intervention group students at the 12-month posttest. After the second year's intervention, however, the two groups reported similar levels.

As predicted, mean perceived difficulty of responding assertively to bullying was lower among intervention group children than among their peers in the control group. A significant Grade  $\times$  Group interaction, shown in Table 6 and Figure 2, was due to a decline in perceived difficulty within the intervention but not the control group.

*Self- and teacher-reported behavior.* Contrary to predictions, the  $F$  values in Table 6 show no group differences in self-reported aggression or victimization. Significant grade slopes show students in both the intervention and control groups reported becoming increasingly aggressive and experiencing decreasing levels of victimization. The means for indirect and direct aggression are combined in Figure 2. Teacher-reported peer interaction skill indicated no group differences.

## Discussion

Many children experiment with dominating others through bullying (Frey et al., in press). Not all become bullies. Whether young people are rewarded for bullying behavior is likely to be a determining factor, hence the need for timely intervention. Despite ample peer support for aggression during pre- and early adolescence (Pellegrini et al., 1999; Pellegrini & Long, 2002; Rodkin, 2006), results of the current study suggest that comprehensive intervention can disrupt the typical increase in bullying (Hanish & Guerra, 2004; Pepler et al., 2002). Participation in the Steps to Respect program for 2 school years was associated with larger declines in problem playground behaviors than participation for 1 year. From pretest levels, rates of bullying and victimization decreased by 19.3% at 6 months and 31.4% by spring of the following year. Similarly, nonbullying aggression decreased by 20% in the first year and by 36.4% in the second. Together, these represent a decline from 66 to 41 aggressive events in an hour on the playground for a class of 25 students. Robust improvements in playground behavior were specifically observed among those involved in bullying-related problems at pretest. Rates for program participants not involved in these behaviors at pretest remained stable across time points, whereas the rates for control group children suggested increases over the school year and across grades. Thus, there is evidence that the program produced both intervention and prevention effects.

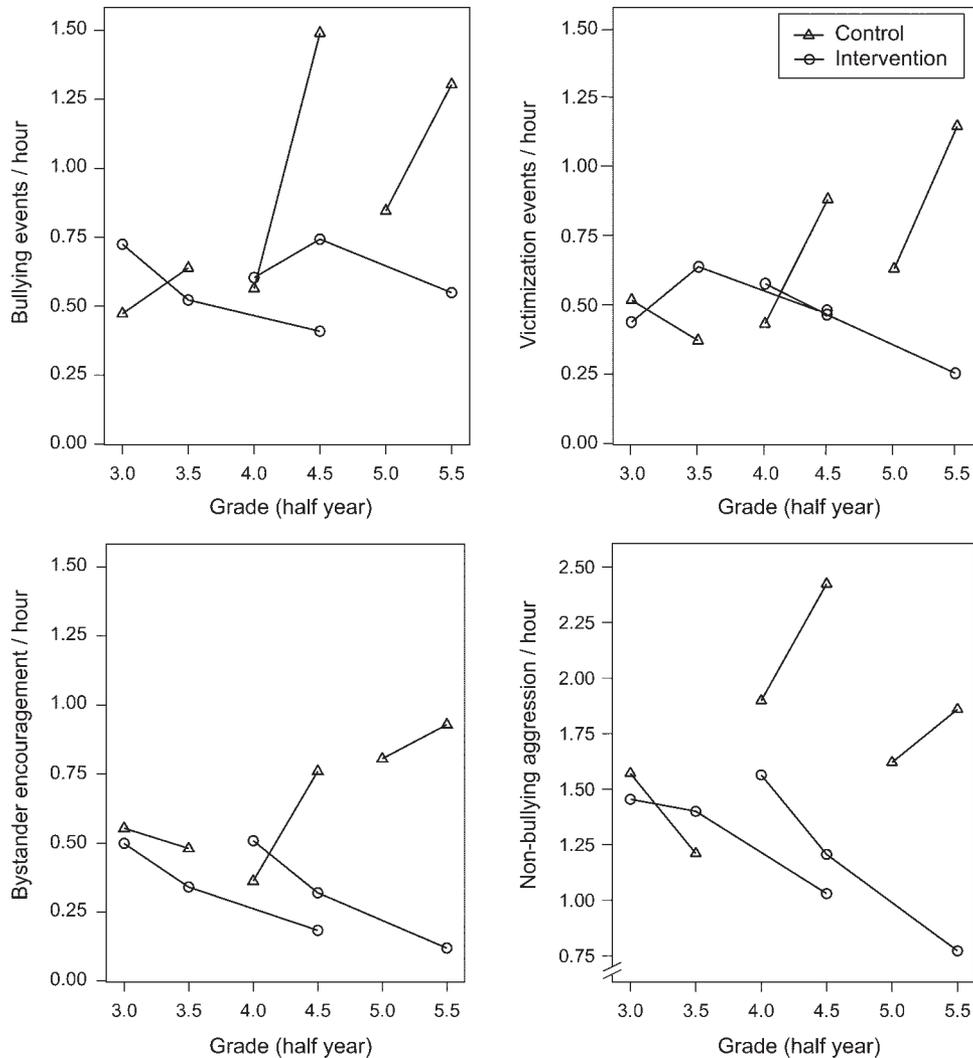


Figure 1. Observed behavior rates by group and six grade points (3.0, fall to 5.5, spring).

### Changes in Observed Bullying and Aggression

After 6 months of intervention, group differences were more robust in bullying than in nonbullying aggression. After 6 months, nonbullying aggression declined only in classrooms of intervention teachers who engaged in more frequent coaching of students involved in bullying (Hirschstein et al., 2007). By 18 months, significant declines in nonbullying aggression appeared within the intervention group and in comparison with the control group. This may indicate that more intensive intervention efforts are required to effect change in nonbullying aggression. One possibility is that our category of nonbullying aggression identified a preponderance of reactive aggression. In contrast to bullying (Sutton, Smith, & Swettenham, 1999), reactive aggression is consistently associated with social-emotional skill deficits (Coe & Dodge, 1998). Addressing skills deficits may require more time and effort than shifting the environmental contingencies thought to influence bullying.

Transactional behavior patterns, in which changes in one behavior contribute to changes in another, may have also played a

role in the pattern of results (Frey & Nolen, in press). Reducing bystander support for aggression was a major program focus. Destructive bystander behavior declined by 40% in the first year and by 72% in the second year. This change may have contributed to second year reductions in bullying behaviors. Reductions in bullying also mean fewer provocations for some students and may foster declines in nonbullying aggression. There is a great need for research that attempts to evaluate such possibilities by explicitly examining the course of change during school interventions.

### Changes in Student Beliefs

Results for two of the four self-report measures of beliefs suggested positive program effects. Program participation appears to have slowed a grade-related increase in the perceived acceptability of bullying and aggression found in the control group and in previous developmental research (Huesmann & Guerra, 1997; Pellegrini & Long, 2002). Also, there were group differences in children's beliefs about the difficulty of responding assertively to

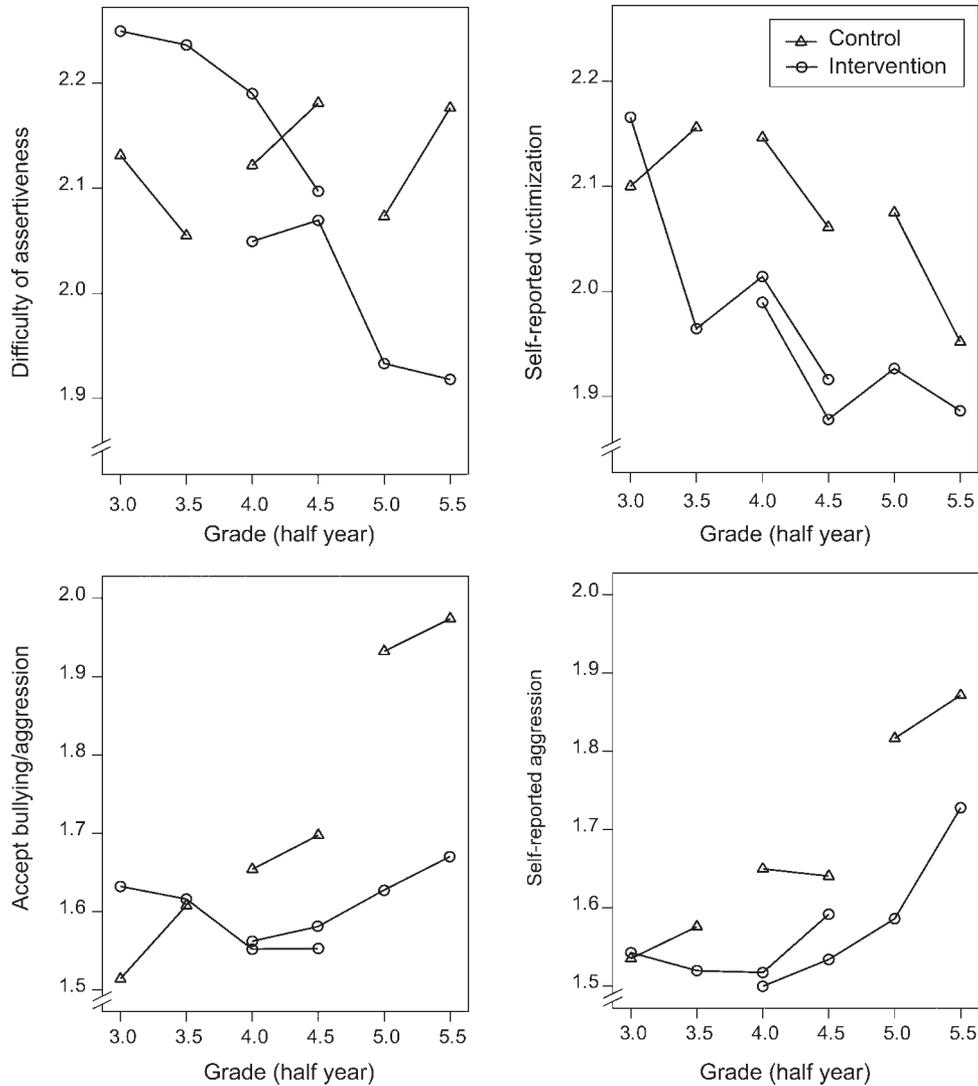


Figure 2. Reported beliefs and behaviors by group and six grade points (3.0, fall to 5.5, spring).

bullying. This suggests that program participants gained confidence in their ability to mount a nonaggressive response, whereas those in the control group did not. Of interest, fifth- and sixth-grade students showed positive effects in their perceived ability to respond assertively after only 6 months of intervention (Frey et al., 2005; Hirschstein et al., 2007). It may require a second year for younger students to feel comfortable responding assertively (“strong and calm”) to bullying.

Given continued adult efforts, it was disturbing to see that longitudinal analyses, in contrast to those of the larger 1-year study (Frey et al., 2005), did not indicate greater perceived adult responsiveness on the part of intervention students. Ratings of responsiveness by program participants declined at the start of the second year, both in comparison to previous ratings and to those of control group students. Perhaps as a result of rising expectations, program participants presumed that the lack of lessons at the start of the year indicated that the program had ended. This supposition is supported by the rebound of ratings, albeit only to the level of the

control group. It is also possible that low internal reliability contributed to the relative volatility of this measure.

Program participants also showed a decrease in perceived personal responsibility. In contrast to results after 6 months, students at the 18-month posttest indicated less willingness to report bullying or remonstrate with friends who bullied. Other studies have found age-related declines in victim defense (Salmivalli & Voeten, 2004). Developmental changes in acceptable responses to bullying may conflict with program effects. Declining perceptions of adult responsiveness may also interfere. Students are less likely to report if they view adults as indifferent (Unnever & Cornell, 2004). As previously noted (Frey et al., 2005), efforts to document adult intervention on the playground were abandoned because of the rarity of the occurrence.

These belief measures encompass both aspects of an adult–child partnership that the Steps to Respect program aims to create in schools. Adults need students to disclose bullying, and children need adults to monitor behavior and provide protection (Frey et al.,

in press). The dynamic nature of student views of adult responsiveness suggests how dependent this partnership is on highly visible adult efforts. Weekly lessons or class meetings provide salient evidence of adult commitment and effort. In light of the low staff-to-child ratios on playgrounds, classroom-based lessons and individual coaching may provide the primary evidence.

### *Self-Perceived Behavior and Experience*

In contrast to observed rates of bullying and victimization, subjective measures provided mixed results. Consistent with other self-reported data (Salmivalli, 2002), older children in both groups viewed themselves as less victimized than younger children. In the intervention group, the decline in perceived victimization was consistent with observed reductions in bullying and victimization and the increased confidence children had in their ability to respond assertively to bullying over time and compared with control students. Paradoxically, control group students also reported grade-related decreases in victimization, despite increased observed victimization rates and no diminution in the perceived difficulty of responding to bullying. Similarly, Salmivalli (2002) found that self-reported, but not peer- or adult-reported, victimization declined from fourth to sixth grade. She found a particular drop in the number of children whose victimization status was based solely on self-report, suggesting that so-called “sensitive” children may apply a more conservative definition of bullying at the end of elementary school than they had earlier.

Surveys in the current study did not include items to screen for truthful responses. We cannot eliminate the possibility that older children become less willing to identify themselves as victims and more willing to identify themselves as potent aggressors. Grade-related increases in the acceptability of aggression, found here and in other studies (Huesmann & Guerra, 1997; Pellegrini & Long, 2002), may contribute to this process. During intermediate elementary grades, boys, in particular, become less willing to seek adult help when victimized (Newman, Murray, & Lussier, 2001), a trend likely to be exacerbated when bullying is perceived to be pervasive (Unnever & Cornell, 2004). Among sixth-grade students, help avoidance is associated with the wish to avoid negative judgments about the self (Ryan, Patrick, & Shim, 2005). As empathy for victims decreases (Rigby & Slee, 1991), identification as a victim may have increasingly negative implications for self-presentation, particularly among boys.

In contrast, aggression can be associated with perceived popularity during this age (Pellegrini et al., 1999; Rodkin et al., 2006; Vaillancourt, Hymel, & McDougall, 2003), and students may become more inclined to report, even celebrate, their transgressions. In the current study, self-reported increases in aggression are consistent with observed increases in the control group and with other research (Hanish & Guerra, 2004). Also in line with previous work (Pellegrini et al., 1999), the perceived acceptability of bullying and aggression in the current study was related to self-reported bullying and aggression at each time point ( $r$ s ranging from .33 to .60), with parallel increases in the two measures across grades. Thus, the program-related changes in bullying behaviors appear in the context of possible developmental changes in the meaning of those behaviors for self-perception and status within the group.

### *Alternative Explanations*

Previous bullying investigators have cautioned against interpreting self-reported behavior as indicative of the actual behaviors of students and adults (O’Connell, Pepler, & Craig, 1999). It is nevertheless possible that the divergence between observed and self-reported behavior reflects the limited time children spent on the playground (about an hour per day). Thus, the intervention may not have produced change outside of the school day, such as on the school bus or in after-school settings. Moreover, we did not measure bullying in the classroom. Other school-based interventions have found significant declines in playground and lunch room aggression, where baseline levels are relatively high, but not in low-frequency venues like the classroom (Grossman et al., 1997). Student self-reports may therefore indicate that the intervention did not yield substantial changes in classrooms, where students spend the most time. The relatively low level of aggression in the classroom may even render episodes there more salient than those on the playground, resulting in a disproportionate impact on student perceptions.

This explanation, however, must then account for the possibility that self-reported victimization is based primarily on events occurring within class, perpetrated by classmates who become more aggressive with each grade (as evidenced by both observations and self-report). Under such circumstances, one would expect an increase in self-reported victimization in the control group rather than a decrease, as we found here.

### *Methodological Considerations*

In contrast to divergent developmental trends in the self-report data, coding by observers blind to condition provided readily interpretable evidence of behavioral changes. Parallel trends in observed bullying and victimization were found in both the intervention and control groups. This is consistent with the reciprocal nature of the behaviors. Bullying requires victims, and those victims are most accessible if they share the same lunch and recess periods, which were grade segregated in these schools.

We are not the first to wonder whether an overreliance on self-reported data in the past has contributed to discouraging intervention results—potentially a problem of limited measurement tools more than inadequate intervention tools (Cornell & Brockenbrough, 2004). We caution that this interpretation of self-report discrepancies may be limited to elementary school. In adolescence, bullying is less confined to school grounds and may be more effectively concealed by sophisticated practitioners (Garandeau & Cillessen, 2006), rendering observations less useful.

In addition, other investigators (Salmivalli, 2002) have argued that self-reports of victimization reflect children’s experience of distress, an important mediator of academic problems (Nishina et al., 2005; Toblin et al., 2005). Discrepancies in self-reported and observed victimization at the end of elementary school suggest fruitful avenues of investigation. Examining the contributions of playground and classroom behavior, normative beliefs, and perceived self-competence in light of the socioecological context (e.g., Espelage et al., 2003; Frey & Nolen, in press; Nishina, Juvonen, & Witkow, 2005) can inform our understanding of underlying developmental processes, as well as provide guidance for measurement choices (Leff, Power, & Goldstein, 2004).

Although behavior observations offer important advantages for program evaluation, costs and time requirements are disincentives

to their use. Even with 50 min of recess observation at each time point, it is likely that we encountered censoring of low-frequency variables like bullying and destructive bystander behavior. That is, subjects who never displayed the event of interest may have done so with sufficiently long observation times. Analyses show that effect sizes are biased downward by censoring in the outcome variable (Stoolmiller, Eddy, & Reid, 2000).

Our conclusions regarding developmental trends must be tempered by limitations in study design. Although the control group encompassed all three grades of the 2-year intervention group, following the control group for 2 years would have increased the accuracy of estimating behavior trends in that group and provided an equal number of survey repetitions.

### Summary

The most robust results of the evaluation came from objective measures of behavior on playgrounds, the arenas in which bullying and victimization are most frequently played out. These showed consistent reductions in problem behavior, especially notable given the increases in bullying typically found at the end of elementary school (Hanish & Guerra, 2004). Comparisons of grade trends in the intervention group with those found in the control group support our conclusion that the program successfully reduced problem behaviors on playgrounds.

By sustaining a visible commitment to a respectful environment, educators in our study conveyed a moral authority that is lacking when socially responsible norms are only given lip service. Reductions in problem behaviors strengthened with a second year of program implementation. The changes observed in destructive bystander behavior were so substantial that the behavior almost disappeared. Although there is considerable theory and evidence that bystanders represent crucial elements in the social-ecological context surrounding bullying, this is the first study to actually show changes in bystander behavior. Removing the support of bystanders signals that bullying is no longer an admired behavior, a potent message for peers to send to each other. Research is needed to test whether such alterations in the environmental contingencies are indeed instrumental in reducing playground bullying.

### References

Andershed, H., Kerr, M., & Stattin, H. (2001). Bullying in school and violence on the streets: Are the same people involved? *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 2, 31–49.

Bandura, A. (1986). *Social foundations of thought and action: A social-cognitive theory*. Englewood, NJ: Prentice Hall.

Boulton, M. J., Trueman, M., Chau, C., Whitehand, C., & Amatya, K. (1999). Concurrent and longitudinal links between friendship and peer victimization: Implications for befriending interventions. *Journal of Adolescence*, 22, 461–466.

Buhs, E. S., Ladd, G. W., & Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Journal of Educational Psychology*, 98, 1–13.

Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development*, 30, 466–480.

Charach, A., Pepler, D., & Ziegler, S. (1995). Bullying at school: A Canadian perspective. *Education Canada*, 35, 12–18.

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Social, emotional, and personality development* (Vol. 3, 5th ed., pp. 779–862). New York: Wiley.

Committee for Children. (2001). *Steps to respect: A bullying prevention program*. Seattle, WA: Author.

Connolly, J., Pepler, D., Craig, W., & Taradash, A. (2000). Dating experiences of bullies in early adolescence. *Child Maltreatment*, 5, 299–310.

Cornell, D. G., & Brockenbrough, K. (2004). Identification of bullies and victims: A comparison of methods. *Journal of School Violence*, 3, 63–87.

Craig, W. M., & Pepler, D. J. (1997). Observations of bullying and victimization in the schoolyard. *Canadian Journal of School Psychology*, 13, 41–59.

Craig, W. M., Pepler, D. J., & Atlas, R. (2000). Observations of bullying on the playground and in the classroom [Special issue]. *School Psychology International*, 21, 22–36.

DeVoe, J. F., & Kaffenberger, S. (2005). *Student reports of bullying: Results from the 2001 school crime supplement to the National Crime Victimization Survey* (NCES Publication No. 2005–310). Washington, DC: U.S. Government Printing Office.

Dodge, K. A., Price, J. M., Coie, J. D., & Christopoulos, C. (1990). On the development of aggressive dyadic relationships in boys' peer groups. *Human Development*, 33, 260–270.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures. *Psychological Methods*, 1, 170–177.

Espelage, D. L., Holt, M. K., & Henkel, R. R. (2003). Examination of peer-group contextual effects on aggression during early adolescence. *Child Development*, 74, 205–220.

Espelage, D. L., & Swearer, S. M. (2003). Research on school bullying and victimization: What have we learned and where do we go from here? *School Psychology Review*, 32, 365–382.

Farmer, T. W., Leung, M.-C., Pearl, R., Rodkin, P. C., Cadwallader, T. W., & Van Acker, R. (2002). Deviant or diverse peer groups? The peer affiliations of aggressive elementary students. *Journal of Educational Psychology*, 94, 611–620.

Frey, K. S., Edstrom, L. V., & Hirschstein, M. K. (in press). School bullying: A crisis or an opportunity? In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *The international handbook of school bullying*. Mahwah, NJ: Erlbaum.

Frey, K. S., Hirschstein, M. K., Snell, J. L., Van Schoiack Edstrom, L., MacKenzie, E. P., & Broderick, C. J. (2005). Reducing playground bullying and supporting beliefs: An experimental trial of the Steps to Respect program. *Developmental Psychology*, 41, 479–490.

Frey, K. S., & Nolen, S. B. (in press). Taking "steps" toward ecological change: A transactional model of school-wide social competence and bullying intervention. In J. Meece & J. Eccles (Eds.), *Schooling effects on children: Theory, methods, and applications*. Mahwah, NJ: Erlbaum.

Garandau, C., & Cillessen, A. H. N. (2005). From indirect to invisible aggression: A conceptual view on bullying and peer group manipulation. *Aggression and Violent Behavior*, 11, 612–625.

Grossman, D. C., Neckerman, H. J., Koepsell, T. D., Liu, P. Y., Asher, K. N., Beland, K., et al. (1997). Effectiveness of a violence prevention curriculum among children in elementary school: A randomized controlled trial. *Journal of the American Medical Association*, 277, 1605–1611.

Hanish, L. D., & Guerra, N. G. (2004). Aggressive victims, passive victims, and bullies: Developmental continuity or developmental change? *Merrill-Palmer Quarterly*, 50, 17–38.

Hanish, L. D., Ryan, P., Martin, C. L., & Fabes, R. A. (2005). The social context of young children's peer victimization. *Social Development*, 14, 2–19.

Hedeker, D., Gibbons, R., & Flay, B. R. (1994). Random-effects regression

- models for clustered data with an example from smoking prevention research. *Journal of Consulting and Clinical Psychology*, 62, 757–765.
- Henry, D., Guerra, N. G., Huesmann, L. R., Tolan, P., Van Acker, R., & Eron, L. (2000). Normative influences on aggression in urban elementary school classrooms. *American Journal of Community Psychology*, 28, 59–81.
- Hirschstein, M. K., Edstrom, L. V., Frey, K. S., Snell, J. L., & MacKenzie, E. P. (2007). Walking the talk in bullying prevention: Teacher implementation variables related to initial impact of the Steps to Respect program. *School Psychology Review*, 36, 3–21.
- Hubbard, J. A., Dodge, K. A., Cillessen, A. H. N., Coie, J. D., & Schwartz, D. (2001). The dyadic nature of social information processing in boys' reactive and proactive aggression. *Journal of Personality and Social Psychology*, 80, 268–280.
- Huesmann, L. R. (1988). An information processing model for the development of aggression. *Aggressive Behavior*, 14, 13–24.
- Huesmann, L. R., & Guerra, N. G. (1997). Children's normative beliefs about aggression and aggressive behavior. *Journal of Personality and Social Psychology*, 72, 408–419.
- Jenson, J. M., & Dietrich, W. A. (2007). Effects of a skilled-based prevention program on bullying and bully victimization among elementary school children. *Prevention Science*, 8, 285–296.
- Juvonen, J., Nishina, A., & Graham, S. (2000). Peer harassment, psychological adjustment, and school functioning in early adolescence. *Journal of Educational Psychology*, 92, 349–359.
- Kallestad, J. H., & Olweus, D. (2003). Predicting teachers' and schools' implementation of the Olweus Bullying Prevention Program: A multi-level study. *Prevention & Treatment*, 6, Article 21. Retrieved March 4, 2004, from <http://journals.apa.org/prevention/volume6>
- Kochenderfer-Ladd, B. (2003). Identification of aggressive and asocial victims and the stability of their peer victimization. *Merrill-Palmer Quarterly*, 49, 401–425.
- Leff, S. S., Power, T. J., & Goldstein, A. B. (2004). Outcome measures to assess the effectiveness of bullying-prevention programs in the schools. In D. L. Espelage & S. M. Swearer (Eds.), *Bullying in American schools: A social-ecological perspective on prevention and intervention* (pp. 269–293). Mahwah, NJ: Erlbaum.
- Musher-Eizenman, D. R., Boxer, P., Danner, S., Dubow, E. F., Goldstein, S. E., & Heretick, D. M. L. (2004). Social-cognitive mediators of the relation of environmental and emotion regulation factors to children's aggression. *Aggressive Behavior*, 30, 389–408.
- Nansel, T. R., Overpeck, M. D., Pilla, R. S., Ruan, J., Simons-Morton, B., & Scheidt, P. C. (2001). Bullying behaviors among U.S. youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association*, 285, 2094–2100.
- Newman, R. S., Murray, B., & Lussier, C. (2001). Confrontation with aggressive peers at school: Students' reluctance to seek help from the teacher. *Journal of Educational Psychology*, 93, 398–410.
- Nishina, A., Juvonen, J., & Witkow, M. R. (2005). Sticks and stones may break my bones, but names will make me feel sick: The psychosocial, somatic, and scholastic consequences of peer harassment. *Journal of Clinical Child and Adolescent Psychology*, 34, 37–48.
- O'Connell, P., Pepler, D., & Craig, W. (1999). Peer involvement in bullying: Insights and challenges for intervention. *Journal of Adolescence*, 22, 437–452.
- Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Oxford, England: Blackwell.
- Olweus, D., Limber, S., & Mihalic, S. (1999). *Blueprints for violence prevention: Bullying prevention program* (Vol. 9). Boulder, CO: Center for the Study and Prevention of Violence, Regents of the University of Colorado.
- Pellegrini, A. D., & Bartini, M. (2000). An empirical comparison of methods of sampling aggression and victimization in school settings. *Journal of Educational Psychology*, 92, 360–366.
- Pellegrini, A. D., Bartini, M., & Brooks, F. (1999). School bullies, victims, and aggressive victims: Factors relating to group affiliation and victimization in early adolescence. *Journal of Educational Psychology*, 91, 216–224.
- Pellegrini, A. D., & Long, J. D. (2002). A longitudinal study of bullying, dominance, and victimization during the transition from primary school through secondary school. *British Journal of Developmental Psychology*, 20, 259–280.
- Pepler, D. J., Craig, W. M., Connolly, J., & Henderson, K. (2002). Aggression and substance use in early adolescence: My friends made me do it. In C. Wekerle & A. M. Wall (Eds.), *The violence and addiction equation: Theoretical and clinical issues in substance abuse and relationship violence* (pp. 153–168). Philadelphia: Brunner/Mazel.
- Pepler, D. J., Craig, W. M., & O'Connell, P. (1999). Understanding bullying from a dynamic systems perspective. In A. Slater & D. Muir (Eds.), *The Blackwell reader in developmental psychology* (pp. 440–451). Malden, MA: Blackwell.
- Rigby, K., & Slee, P. T. (1991). Bullying among Australian school children: Reported behavior and attitudes toward victims. *Journal of Social Psychology*, 131, 615–627.
- Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2006). They're cool: Social status and peer group supports for aggressive boys and girls. *Social Development*, 15, 175–204.
- Ryan, A. M., Patrick, H., & Shim, S.-O. (2005). Differential profiles of students identified by their teacher as having avoidant, appropriate, or dependent help-seeking tendencies in the classroom. *Journal of Educational Psychology*, 97, 275–285.
- Salmivalli, C. (1999). Participant role approach to bullying: Implications for intervention. *Journal of Adolescence*, 22, 453–459.
- Salmivalli, C. (2002). Is there an age decline in victimization by peers at school? *Educational Research*, 44, 269–277.
- Salmivalli, C., & Helteenvuori, T. (2007). Reactive, but not proactive aggression predicts victimization among boys. *Aggressive Behavior*, 33, 198–206.
- Salmivalli, C., & Voeten, M. (2004). Connections between attitudes, group norms, and behavior and bullying situations. *International Journal of Behavior Development*, 28, 246–258.
- Sameroff, A. J., & MacKenzie, M. J. (2003). Research strategies for capturing transactional models of development: The limits of the possible. *Development and Psychopathology*, 15, 613–640.
- Schafer, M., Werner, N. E., & Crick, N. R. (2002). A comparison of two approaches to the study of negative peer treatment: General victimization and bully-victim problems among German schoolchildren. *British Journal of Developmental Psychology*, 20, 281–306.
- Schwartz, D. (2000). Subtypes of victims and aggressors in children's play groups. *Journal of Abnormal Child Psychology*, 28, 181–192.
- Schwartz, D., Gorman, A. H., Nakamoto, J., & Toblin, T. M. (2005). Victimization in the peer group and children's academic functioning. *Journal of Educational Psychology*, 97, 425–435.
- Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis*. New York: Oxford Press.
- Smith, J. D., Schneider, B. H., Smith, P. K., & Ananiadou, K. (2004). The effectiveness of whole-school antibullying programs: A synthesis of evaluation research. *School Psychology Review*, 33, 547–560.
- Smith-Khuri, E., Iachan, R., Scheidt, P. C., Overpeck, M. D., Gabhainn, S. N., Pickett, W., et al. (2004). A cross-national study of violence-related behaviors in adolescents. *Archives of Pediatric and Adolescent Medicine*, 158, 539–544.
- Snell, J. L., & Frey, K. S. (2000). *Playground aggressors, targets and bystanders (PATAB)*. Unpublished manuscript, Committee for Children, Seattle, WA.
- Snyder, J., Reid, J., Stoolmiller, M., Howe, G., Brown, H., Dagne, G., et al. (2006). The role of behavior observations in measurement systems for randomized prevention trials. *Prevention Science*, 7, 43–56.
- Stoolmiller, M., Eddy, J. M., & Reid, J. B. (2000). Detecting and describ-

## PROGRAM-RELATED DECLINES IN OBSERVED BULLYING

- ing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology*, 68, 296–306.
- Sutton, J., Smith, P. K., & Swettenham, J. (1999). Bullying and “theory of mind”: A critique of the “social skills deficit” view of anti-social behavior. *Social Development*, 8, 117–127.
- Swearer, S. M., & Espelage, D. L. (2004). A social–ecological framework of bullying among youth. In D. L. Espelage & S. M. Swearer (Eds.), *Bullying in American schools: A social–ecological perspective on prevention and intervention* (pp. 1–12). Mahwah, NJ: Erlbaum.
- Terasahjo, T., & Salmivalli, C. (2003). “She is not actually bullied:” The discourse of harassment in student groups. *Aggressive Behavior*, 29, 134–154.
- Toblin, R. L., Schwartz, D., Gorman, A. H., & Abou-ezzedine, T. (2005). Social–cognitive and behavioral attributes of aggressive victims of bullying. *Journal of Applied Developmental Psychology*, 26, 329–346.
- Unnever, J. D., & Cornell, D. G. (2003). The culture of bullying in middle school. *Journal of School Violence*, 2, 5–27.
- Unnever, J. D., & Cornell, D. G. (2004). Middle school victims of bullying: Who reports being bullied? *Aggressive Behavior*, 30, 373–388.
- Vaillancourt, T., Hymel, S., & McDougall, P. (2003). Bullying is power: Implications for school-based intervention. *Journal of Applied School Psychology*, 19, 157–176.
- Walker, H. M., & McConnell, S. R. (1995). *Walker–McConnell Scale of Social Competence and School Adjustment: Elementary version*. San Diego, CA: Singular.

|

|